

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:12:44 ; Search time 200 Seconds
(without alignments)
2926.263 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 M8A6LGRWLLWALLRLAQ.....GSIKAGALSNFLPNNNGROF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

Database :

A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	8	ADJ27297
2	6909	100.0	1332	9	ADY60740 Human NPC
3	6909	100.0	1332	9	AEB93570 Human NPC
4	6896	99.8	1332	4	AAM79169 Human pro
5	6896	99.8	1332	4	AAG65638 SSD-conta
6	6872.5	99.5	1359	4	AAM79168 Human pro
7	6872.5	99.5	1359	8	ADJ27337 Human NPC
8	6872.5	99.5	1359	9	ADY60780 Human NPC
9	6872.5	99.5	1359	9	AEB93610 Human NPC
10	6536	94.6	1344	4	AEG22693 Novel hum
11	5421.5	78.5	1331	8	ADJ27295 Rat NPCIL
12	5421.5	78.5	1331	9	ADY60738 Rat Niema
13	5421.5	78.5	1331	9	AEB93568 Rat NPCIL
14	5407	78.3	1333	8	ADJ27305 Mouse NPC
15	5407	78.3	1333	9	ADY60748 Mouse Nie
16	5407	78.3	1333	9	AEB93578 Mouse NPC
17	4466	64.6	982	4	ABG22691 Novel hum
18	2402.5	34.8	1278	3	AAB88445 Human NPC
19	2402.5	34.8	1278	3	AAB842983 Human ORF
20	2402.5	34.8	1278	8	ADJ39879 Human myo
21	2402.5	34.8	1278	8	ADU06723 Novel NPC
22	2385	34.5	1319	2	AAM88446 Mouse NPC
23	2262	32.7	1287	4	AB61737 Drosophil
24	2262	32.7	1287	8	ADS96670 Drosophil

ALIGNMENTS

RESULT 1

ADJ27297

ID ADJ27297 standard; protein; 1332 AA.

XX AC ADJ27297;

XX DT 20-MAY-2004 (first entry)

XX DE Human NPCIL1.

XX KW Niemann-Pick disease; type C1; gene-like 1; NPCIL1; trans-golgi network;

XX KW plasma membrane; transport signal; promoter;

XX KW sterol regulated element binding protein 1; SREBP1;

XX KW binding consensus sequence; transmembrane domain; sterol-sensing domain;

XX KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;

XX KW cholesterol absorption; serum cholesterol; hyperlipidaemia;

XX KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX OS Homo sapiens.

XX PN WO2004009772-A2.

XX PD 29-JAN-2004.

XX PF 17-JUL-2003; 2003WO-US022467.

XX PR 19-JUL-2002; 2002US-0397442P.

XX PI (SCHE) SCHERING CORP.

XX PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX DR WPI; 2004-132945/13.

XX DR N-PSDB; ADJ27296.

XX XX GENBANK; AF192522.

PT New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting or identifying antagonists of NPCIL1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

PS Claim 11; SEQ ID NO 4; 125pp; English.

XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1 (NPCIL1) polypeptide. NPCIL1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and

Abb58629 Drosophil
Aaw88447 Yeast NPC
Aaw88448 Caenorhab
Aam80153 Human pro
Aam80152 Human pro
Aag65637 SSD domal
Aar75375 Human pat
Aaw52200 Human pat
Aaw72969 Human pat
Aab67163 Human pat
Aae19830 Human pat
Abj10931 TRC8 rela
Aag79571 Human pat
Abu62275 Human pat
Add46678 Human pro
Ade94224 Human ptc
Adh62731 Human pat
Ade48989 Human pat
Aaw52199 Mouse pat
Aaw72968 Mouse pat
Aab67159 Murine pa

25 1864.5 27.0 1223 4 ABB58629
26 1329 19.2 1170 2 AAW88447
27 1078 15.6 1296 2 AAW88448
28 963 13.9 229 4 AAM80153
29 963 13.9 229 4 AAM80152
30 942 13.6 194 4 AAG65637
31 651.5 9.4 1447 2 AAR75375
32 651.5 9.4 1447 2 AAW52200
33 651.5 9.4 1447 2 AAW72969
34 651.5 9.4 1447 4 AAB67163
35 651.5 9.4 1447 5 AAE19830
36 651.5 9.4 1447 5 ABJ10931
37 651.5 9.4 1447 5 AAG79571
38 651.5 9.4 1447 7 ABU62275
39 651.5 9.4 1447 7 ADD46678
40 651.5 9.4 1447 7 ADE94224
41 651.5 9.4 1447 7 ADH62731
42 651.5 9.4 1447 8 ADE48989
43 646.5 9.4 1434 2 AAW52199
44 646.5 9.4 1434 2 AAW72968
45 646.5 9.4 1434 4 AAB67159

CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a
CC sterol regulated element binding protein 1 (SREBP1) binding consensus
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
CC sensing domain (SSD) which is involved in sensing cholesterol levels,
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1
CC has 42 amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.

XX
SQ Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 8; Length 1332;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSPPYTHIOPGYCAFYDECGKPELSSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSPPYTHIOPGYCAFYDECGKPELSSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDVDSRVRVPA 180
DB 121 VNLHCHNTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDVDSRVRVPA 180
QY 181 ATLAVTGTCGVGSAACNARWLNFGDGTGNGLAPLDITPHLEPGQAVGSGIQLNEGV 240
DB 181 ATLAVTGTCGVGSAACNARWLNFGDGTGNGLAPLDITPHLEPGQAVGSGIQLNEGV 240
QY 241 ARCNEQSGDDVATCSQDCAASPAIARPAQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNEQSGDDVATCSQDCAASPAIARPAQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKVMDPKGTSLSDKLSFSFHTLLGQFFQCGWGTWVASWELTILVLSV 360
DB 301 LLVGFVAPARDKSKVMDPKGTSLSDKLSFSFHTLLGQFFQCGWGTWVASWELTILVLSV 360
QY 361 IPVVALAAGLVFELTTPVELMSAPNSQARSKAFHDOHFGPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFELTTPVELMSAPNSQARSKAFHDOHFGPFRTNQVILTAPNRSY 420
QY 421 RYDSSLILGPKNFGSILDLILLLELLELRLHQLVMSPEAQARNISLODICYAPLNPDT 480
DB 421 RYDSSLILGPKNFGSILDLILLLELLELRLHQLVMSPEAQARNISLODICYAPLNPDT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGFTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCWADYGAPVFPFLATGGYKGDYSABEALIMTFSLNYPAGDPRLAQAKLMEAFLEEM 600
DB 541 SCWADYGAPVFPFLATGGYKGDYSABEALIMTFSLNYPAGDPRLAQAKLMEAFLEEM 600
QY 601 RAFQRMAGNFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGYSYSSWSRV 660
DB 601 RAFQRMAGNFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGYSYSSWSRV 660
QY 661 MVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSILVQVFPFLVLSVAGDNIFFVLE 720
DB 661 MVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSILVQVFPFLVLSVAGDNIFFVLE 720
QY 721 YQRLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTMPAVRTFALTSLGLAV 780
DB 721 YQRLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEARSLDVCCVKPQELPPPGQEGLLGFFQKAYAPFL 840

DB 781 ILDFLLQMSAFVALLSLDSKQEARSLDVCCVKPQELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHMTRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHMTRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTGLYFNSEAGMNAICSSAGCANNFSTQKIYATEFPPEQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTGLYFNSEAGMNAICSSAGCANNFSTQKIYATEFPPEQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNKMSTITGVSVRPSYEQPHKYLFWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNKMSTITGVSVRPSYEQPHKYLFWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQVLAASFMAHYKPKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTVNLTSQVLAASFMAHYKPKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEVFPYTIITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEVFPYTIITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSTVMTLVDTGFMALWDISYNAVSLINLVSAVMSVEFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSTVMTLVDTGFMALWDISYNAVSLINLVSAVMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFRNLNLTLLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFRNLNLTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEEAAVAVMVASCPNHPSPRVSTADNTIYVNHSPSGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEEAAVAVMVASCPNHPSPRVSTADNTIYVNHSPSGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332
RESULT 2
ADY60740
ID ADY60740 standard; protein; 1332 AA.
XX
AC ADY60740;
XT 19-MAY-2005 (first entry)
XX Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 4.
XX Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX Homo sapiens.
XX WO2005015988-A1.
XX 24-FEB-2005.
XX 16-DEC-2003; 2003WO-US040113.
XX 17-JUL-2003; 2003US-00621758.
XX 22-AUG-2003; 2003US-00646301.
XX 16-SEP-2003; 2003US-00663208.
XX (SCHE) SCHERING CORP.
XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2005-284403/29.
XX N-PSDB; ADY60739, ADY60787.
DR

XX Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
 PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
 PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
 XX
 PS Claim 30; SEQ ID NO 4; 146pp; English.
 XX

CC The invention relates to an isolated Niemann-Pick C1-like protein 1
 CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
 CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,
 CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
 CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
 CC recombinant vector comprising the polynucleotide, a host cell comprising
 CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
 CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
 CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
 CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
 CC produce any functional NPC1L1 protein), an offspring or progeny of the
 CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
 CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
 CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
 CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
 CC dosage form, and information indicating that NPC1L1 is a target of
 CC ezetimibe or substituted azetidinone), decreasing the level of intestinal
 CC sterol or Salpha-sterol absorption in a subject (involving reducing the
 CC level of expression of NPC1L1 in the subject), identifying an antagonist
 CC of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a
 CC functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist
 CC of NPC1L1 which is useful for inhibiting or decreasing the level of
 CC NPC1L1 mediated sterol or Salpha-sterol uptake in a subject such as
 CC mouse, rat or human. The level of expression of NPC1L1 in the subject is
 CC reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is
 CC useful for screening a sample for intestinal sterol or Salpha-sterol
 CC absorption antagonist. The NPC1L1 antagonists identified are useful for
 CC treating medical conditions caused or mediated by NPC1L1, e.g.,
 CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
 CC arteriosclerosis. The present sequence represents a mammalian NPC1L1
 CC protein.
 XX
 SQ Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 9; Length 1332;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
 DB 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
 QY 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVLSLEASLSITKALLTRCPACSDNF 120
 DB 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVLSLEASLSITKALLTRCPACSDNF 120
 QY 121 VNLHCHNTCSNQSIFINTRVAQIAGAGLPVAVYAEAPYQHSFAEQSDSCSRVRPAA 180
 DB 121 VNLHCHNTCSNQSIFINTRVAQIAGAGLPVAVYAEAPYQHSFAEQSDSCSRVRPAA 180
 QY 181 ATLVGTMCGVYGSALCNAQRLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
 DB 181 ATLVGTMCGVYGSALCNAQRLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
 QY 241 ARCNSQGDVATCSQDCQCAAPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
 DB 241 ARCNSQGDVATCSQDCQCAAPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
 QY 301 LLVGRVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFGCGWGTWASWPLTILVLSV 360
 DB 301 LLVGRVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFGCGWGTWASWPLTILVLSV 360
 QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDOHFGFPFRFTNQVILTANRSSY 420
 DB 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDOHFGFPFRFTNQVILTANRSSY 420

QY 421 RYDSELLGPKNPSGILDLDLLELELEQLERLRLHQLVMSPEAQRTSLQDICYAPLNPDNT 480
 DB 421 RYDSELLGPKNPSGILDLDLLELELEQLERLRLHQLVMSPEAQRTSLQDICYAPLNPDNT 480
 QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTKDGTALAL 540
 DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTKDGTALAL 540
 QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLEAEAELEEM 600
 DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLEAEAELEEM 600
 QY 601 RAFORMMAGMFOVTTAERSLEDEINRTAEDLPFATSYIVIFYISIALSGSYSSWSRV 660
 DB 601 RAFORMMAGMFOVTTAERSLEDEINRTAEDLPFATSYIVIFYISIALSGSYSSWSRV 660
 QY 661 WYDSKATLGLGGVAVVLGAVMAAMGPPSVGLTRSSILVILQVVPFLVLSVGAONIFIVLE 720
 DB 661 WYDSKATLGLGGVAVVLGAVMAAMGPPSVGLTRSSILVILQVVPFLVLSVGAONIFIVLE 720
 QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICPFLGALTMPAVRTALTSLGLAV 780
 DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICPFLGALTMPAVRTALTSLGLAV 780
 QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCCKPQELPPPGQEGLLGLFPQKAYAPFL 840
 DB 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCCKPQELPPPGQEGLLGLFPQKAYAPFL 840
 QY 841 LHMITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAVP 900
 DB 841 LHMITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAVP 900
 QY 901 YFVVTILGYNFSSEAGNNAICSSAGCNVNSFTQKIQVATEFPQSYLAIPASSWDDPIDW 960
 DB 901 YFVVTILGYNFSSEAGNNAICSSAGCNVNSFTQKIQVATEFPQSYLAIPASSWDDPIDW 960
 QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLFWFLNDRP 1020
 DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLFWFLNDRP 1020
 QY 1021 NIKCPKGLAAYSTVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
 DB 1021 NIKCPKGLAAYSTVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
 QY 1081 RKVPGTDPAFEPVPTITNVFYEQYLTILPEGLFMLSCLVCLVTEAVSCLLGLDLRSGLL 1140
 DB 1081 RKVPGTDPAFEPVPTITNVFYEQYLTILPEGLFMLSCLVCLVTEAVSCLLGLDLRSGLL 1140
 QY 1141 NLLSIVMILVDVTGPMALWDISYNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER 1200
 DB 1141 NLLSIVMILVDVTGPMALWDISYNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER 1200
 QY 1201 AKEATISMGSAVPAVAMTNLPGIIVLGLAKAQLIQIPFRLNLLITLLGLHLGLVFLPV 1260
 DB 1201 AKEATISMGSAVPAVAMTNLPGIIVLGLAKAQLIQIPFRLNLLITLLGLHLGLVFLPV 1260
 QY 1261 ILSYVGPDPVPALALEQKRAEBAVAAVMVASCPNHPSPRVSTADNIYVNHSPGSGIKGACA 1320
 DB 1261 ILSYVGPDPVPALALEQKRAEBAVAAVMVASCPNHPSPRVSTADNIYVNHSPGSGIKGACA 1320
 QY 1321 ISNPLPNNGRQP 1332
 DB 1321 ISNPLPNNGRQP 1332
 QY 1321 ISNPLPNNGRQP 1332
 DB 1321 ISNPLPNNGRQP 1332
 RESULT 3
 ID AEB93570
 XX AEB93570 standard; protein; 1332 AA.
 AC AEB93570;
 XX
 XX 06-OCT-2005 (first entry)
 XX

DE Human NPC1l1 protein.
XX Protein engineering; NPC1l1; antidiabetic; antiarteriosclerotic;
KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
KW arteriosclerosis; human.
XX Homo sapiens.
OS
XX WO2005069900-A2.
PN
XX
XX 04-AUG-2005.
PD
XX
XX 14-JAN-2005; 2005WO-US001469.
PF
XX
XX 16-JAN-2004; 2004US-0537341P.
PR
XX (MERI) MERCK & CO INC.
PA
XX Garcia-Calvo M;
PI
XX WPI; 2005-564070/57.
DR N-PSDB; AEB93569.
DR
XX
XX Identifying ligand of NPC1l1 for stimulating the activity of NPC1l1, by
PT contacting NPC1l1 with detectably labeled substituted 2-azetidinone
PT glucuronide and a candidate compound and determining if compound binds to
PT human NPC1l1.
XX
XX Example 2; SEQ ID NO 4; 215pp; English.
PS
XX
XX The invention relates to identifying a ligand of NPC1l1. The method
CC involves contacting human NPC1l1 with a detectably labeled substituted 2-
CC azetidinone glucuronide and a candidate compound and determining whether
CC the candidate compound binds to human NPC1l1. In identifying a ligand of
CC NPC1l1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10
CC nM or lower. The detectably labeled substituted 2-azetidinone glucuronide
CC is 35 S-labeled compound 2, given in the specification. NPC1l1 ligands
CC are useful for stimulating or blocking the activity of NPC1l1, and for
CC treating conditions caused or mediated by NPC1l1. It is useful for
CC reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart
CC disease, stroke, or arteriosclerosis. The present sequence represents a
CC human NPC1l1, a N-glycosylated protein.
XX
SQ Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 9; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHHPGYCAFYDECGKPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHHPGYCAFYDECGKPELSGLMTLSNVSCLSN 60

QY 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVLEASLSITKALITRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVLEASLSITKALITRCPACSDNF 120

QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFABQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFABQSDSCSRVRPAA 180

QY 181 ATLAVTGTCGYGSAICNAQRWLNFGQDTGNGLAPLDTIFHLAEPGQAVGSGIQPLNEGV 240
DB 181 ATLAVTGTCGYGSAICNAQRWLNFGQDTGNGLAPLDTIFHLAEPGQAVGSGIQPLNEGV 240

QY 241 ARCNESQDDVATCSQDCQCAAPARPAQALDSTFYLQMPGSLVLIITLCSVFPAVTTI 300
DB 241 ARCNESQDDVATCSQDCQCAAPARPAQALDSTFYLQMPGSLVLIITLCSVFPAVTTI 300

QY 301 LLVGFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFPQCGTGWASWPLTILVLSV 360
DB 301 LLVGFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFPQCGTGWASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTTELTTDPVELMSAPNSQARSEKAFHDQHFQFPFRNQVILTPAPRSSY 420
DB 361 IPVVALAAGLVFTTELTTDPVELMSAPNSQARSEKAFHDQHFQFPFRNQVILTPAPRSSY 420

QY 421 RYDSLLLGPKNFSGILDLDLLELLELQERLRLHQLQWSPQAQRNLSLODICYAPLNPONT 480
DB 421 RYDSLLLGPKNFSGILDLDLLELLELQERLRLHQLQWSPQAQRNLSLODICYAPLNPONT 480

QY 481 SLYDCCINSILQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL 540
DB 481 SLYDCCINSILQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL 540

QY 541 SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTSLNNYPAGDPRLAQAKLWEBAFLBEM 600
DB 541 SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTSLNNYPAGDPRLAQAKLWEBAFLBEM 600

QY 601 RAFQRMAGMFQVTTAERSLEDEINRTTABDLPIFATSYIVIFLYISLALGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTTAERSLEDEINRTTABDLPIFATSYIVIFLYISLALGSYSSWSRV 660

QY 661 MVDKATILGLGGVAVVLGAVMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLE 720
DB 661 MVDKATILGLGGVAVVLGAVMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLE 720

QY 721 YQRLPRRPGEPREVIHIGRALGRVAPSMMLCSLSEAI CFFLGALTPMPAVRFTALTSLAV 780
DB 721 YQRLPRRPGEPREVIHIGRALGRVAPSMMLCSLSEAI CFFLGALTPMPAVRFTALTSLAV 780

QY 781 ILDFLLQMSAFVALLSDSKROBASRLDVCCKVPOELPPPGQEGELLGFKQAYAPFL 840
DB 781 ILDFLLQMSAFVALLSDSKROBASRLDVCCKVPOELPPPGQEGELLGFKQAYAPFL 840

QY 841 LHWITRGVVLVLLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
DB 841 LHWITRGVVLVLLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900

QY 901 YFVTTLYGNFSEAGMNAICSSAGCNSFTQKIYATEPPEQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTLYGNFSEAGMNAICSSAGCNSFTQKIYATEPPEQSYLAIPASSWVDDFIDW 960

QY 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNKMSITMGSVRPSVEQPHKYLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNKMSITMGSVRPSVEQPHKYLWFLNDRP 1020

QY 1021 NIKCPKGLAAYSTVNLTSQVLSRPMAYHKPKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTVNLTSQVLSRPMAYHKPKNSQDYTEALRAARELAANITADL 1080

QY 1081 RKVPGTDPAFEPVPTIITNVFYEYLTILPEGLFMLSICLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEPVPTIITNVFYEYLTILPEGLFMLSICLVPTFAVSCLLGLDLRSGLL 1140

QY 1141 NLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200

QY 1201 AKEATISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIFFRNLMLITLLGLHLGLVFLPV 1260
DB 1201 AKEATISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIFFRNLMLITLLGLHLGLVFLPV 1260

QY 1261 ILSYVGPDPVNPALAEQKAEBAVAAVMVASCPNHPSRVSTADNIYVNHSPFEGSIKAGA 1320
DB 1261 ILSYVGPDPVNPALAEQKAEBAVAAVMVASCPNHPSRVSTADNIYVNHSPFEGSIKAGA 1320

QY 1321 ISNFLPNNRGQF 1332
DB 1321 ISNFLPNNRGQF 1332

RESULT 4
AAW79169
ID AAW79169 standard; protein; 1332 AA.
XX

AC	AAM79169;	
XX	06-NOV-2001 (first entry)	
DE	Human protein SEQ ID NO 1831.	
XX		
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157190-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	05-FEB-2001; 2001WO-US004098.	
XX		
PR	03-FEB-2000; 2000US-00496914.	
PR	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
PI	Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX		
DR	WPI; 2001-476283/51.	
DR	N-PSDB; AAK52302.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT	in diagnosis and gene therapy.	
XX		
PS	Claim 20; Page 4216-4219; 6221pp; English.	
XX		
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the	
CC	encoded polypeptides (AAM79323-AAK80302) that exhibit activity elating to	
CC	cytokine, cell proliferation or cell differentiation or which may induce	
CC	production of other cytokines in other cell populations. The	
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or	
CC	peptide therapy. The polypeptides have various cytokine-like activities,	
CC	e.g. stem cell growth factor activity, haematopoiesis regulating	
CC	activity, tissue growth factor activity, immunomodulatory activity and	
CC	activin/inhibin activity and may be useful in the diagnosis and/or	
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and	
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111	
CC	(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the	
CC	sequence listing were missing at the time of publication	
XX		
SQ	Sequence 1332 AA;	
	Query Match 99.8%; Score 6896; DB 4; Length 1332;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60	
DB	1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60	
QY	61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKOLVSLASLSITKALLTRCPACSDNF 120	
DB	61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKOLVSLASLSITKALLTRCPACSDNF 120	
QY	121 VNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFABQSDSCSRVRPAA 180	
DB	121 VNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFABQSDSCSRVRPAA 180	

QY	181 ATLA VGTMC GVGYSALCNAQRWLNFGQDGTGNGLA PLDITFHLLEPGQAVSGGIQPLNEGV 240	
DB	181 ATLA VGTMC GVGYSALCNAQRWLNFGQDGTGNGLA PLDITFHLLEPGQAVSGGIQPLNEGV 240	
QY	241 ARCNSQGGDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300	
DB	241 ARCNSQGGDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300	
QY	301 LLVGRVAPARDKSKWDPKKGTSLSDKLSPSTHTLLGQFFQCGWGTVWASWPLTILVLSV 360	
DB	301 LLVGRVAPARDKSKWDPKKGTSLSDKLSPSTHTLLGQFFQCGWGTVWASWPLTILVLSV 360	
QY	361 IPVVALAAGLVFTELTTPDELWSPNSQARSEKAFHQHFGPPFRFTNQTIVITAPNRSY 420	
DB	361 IPVVALAAGLVFTELTTPDELWSPNSQARSEKAFHQHFGPPFRFTNQTIVITAPNRSY 420	
QY	421 RYDSLLLGPKNPSGILDLDLLELLELQERLHLQVWSPQARNISLQDICYAPLNPDT 480	
DB	421 RYDSLLLGPKNPSGILDLDLLELLELQERLHLQVWSPQARNISLQDICYAPLNPDT 480	
QY	481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDMKDHFLYCANAPLTPKDGTTALAL 540	
DB	481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDMKDHFLYCANAPLTPKDGTTALAL 540	
QY	541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLBEM 600	
DB	541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLBEM 600	
QY	601 RAPQRMAGMQVTTAERSLEDEINRTAEDLPITATSYIVIFLYISLALSGYSWSRV 660	
DB	601 RAPQRMAGMQVTTAERSLEDEINRTAEDLPITATSYIVIFLYISLALSGYSWSRV 660	
QY	661 MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720	
DB	661 MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720	
QY	721 YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEAI CFFLGLALTTPMPAVRTALTSGLA V 780	
DB	721 YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEAI CFFLGLALTTPMPAVRTALTSGLA V 780	
QY	781 ILDFLLQMSAFVALLSLSKQESRLDVCCKVPQELPPPGQGGGLLPGFKAYAPPL 840	
DB	781 ILDFLLQMSAFVALLSLSKQESRLDVCCKVPQELPPPGQGGGLLPGFKAYAPPL 840	
QY	841 LHMITRGVVLTLFLALFGVSLYSMCHISVGLDQELALPKDSVILIDYFLFLNRYPFVGAPV 900	
DB	841 LHMITRGVVLTLFLALFGVSLYSMCHISVGLDQELALPKDSVILIDYFLFLNRYPFVGAPV 900	
QY	901 YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATEFPPEQSYLAIPASSWVDDFDW 960	
DB	901 YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATEFPPEQSYLAIPASSWVDDFDW 960	
QY	961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHMSITMGSVRPSVEQFHKLVPFLNDRP 1020	
DB	961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHMSITMGSVRPSVEQFHKLVPFLNDRP 1020	
QY	1021 NIKCPKGLAAYSTSVNLTSDQVLSRPMAYHKLKNSQDYTEALRAARELANITADL 1080	
DB	1021 NIKCPKGLAAYSTSVNLTSDQVLSRPMAYHKLKNSQDYTEALRAARELANITADL 1080	
QY	1081 RKVPGTDPAPFVPPTITNVFYEQVLTLPGLFMLSCLVPTTAVSCLLGLDLRSGLL 1140	
DB	1081 RKVPGTDPAPFVPPTITNVFYEQVLTLPGLFMLSCLVPTTAVSCLLGLDLRSGLL 1140	
QY	1141 NLLSIVMLVDTVGFMALWDISYNVAVSLINIVSAGMSVEFVSHITRSPFAISTKPTWLER 1200	
DB	1141 NLLSIVMLVDTVGFMALWDISYNVAVSLINIVSAGMSVEFVSHITRSPFAISTKPTWLER 1200	
QY	1201 AKEATISNGSAVPAVAVMTNLPGLIVLGLAKAQLIQIOPFFRLNLLITLLGLHLGVLPV 1260	
DB	1201 AKEATISNGSAVPAVAVMTNLPGLIVLGLAKAQLIQIOPFFRLNLLITLLGLHLGVLPV 1260	

Qy	1261	ILSVGPDVNPALALEQKRAEEAAVAVVASCNHPHSRVSTADNIYVNHSPFEGSIKAGA	1320
Db	1261	ILSVGPDVNPALALEQKRAEEAAVAVVASCNHPHSRVSTADNIYVNHSPFEGSIKAGA	1320
Qy	1321	ISNPLPNNGRQF	1332
Db	1321	ISNPLPNNGRQF	1332
RESULT 5			
ID	AAG65638	standard; protein; 1332 AA.	
XX			
AC	AAG65638;		
XX			
DT	07-JAN-2002	(first entry)	
DE			
XX		SSD-containing SSP1 protein sequence.	
KW		SSD; sterol-sensing domain; human; liver; testis; brain; cancer;	
KW		neotropic; neuroprotective; antidiabetic; antiarteriosclerotic;	
KW		cytostatic; antilipemic; SSP1.	
XX			
OS		Homo sapiens.	
XX			
PN	WO200170974-A1.		
XX			
PD	27-SEP-2001.		
XX			
PF	22-MAR-2001; 2001WO-JP002279.		
XX			
PR	24-MAR-2000; 2000JP-00088595.		
XX			
PA	(TAKE) TAKEDA CHEM IND LTD.		
XX			
PI	Taniyama Y, Kita S, Komiyama T;		
XX			
DR	WPI; 2001-611501/70.		
DR	N-PSDB; AAI66914.		
XX			
PT	New sterol-sensing domain-containing protein for diagnosing and screening		
PT	candidate compounds in drug development for diabetes, obesity, cancer,		
PT	arteriosclerosis, hyperlipidemia and neurodegenerative disorders.		
XX			
PS	Claim 7; Page 115-122; 171pp; Japanese.		
XX			
CC	The invention provides a novel SSD (sterol-sensing domain)-containing		
CC	protein. The protein originates from human liver, human testis or human		
CC	brain. The protein can be expressed by standard recombinant methodology.		
CC	The proteins, encoded DNAs and antibodies are useful in diagnosis and		
CC	screening candidate compounds in drug development for diabetes, obesity,		
CC	cancer, arteriosclerosis, hyperlipemia, neurodegenerative disorders such		
CC	as Alzheimer's disease and neural disorders. The present sequence		
CC	represents the SSP1 protein which contains the SSD domain		
XX			
SQ	Sequence 1332 AA;		
Query Match 99.8%; Score 6896; DB 4; Length 1332;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MAEAGRGWLLWALLLLRLAQSEPYTTTHQPGYCAFYDECCKNPGLSGMTLSNVSCLSN	60
Db	1	MAEAGRGWLLWALLLLRLAQSEPYTTTHQPGYCAFYDECCKNPGLSGMTLSNVSCLSN	60
Qy	61	TPARKITGDHILLOKICPRLYTCPTNTQACCSAKQLVLEASLITKALLTRCPACSDNF	120
Db	61	TPARKITGDHILLOKICPRLYTCPTNTQACCSAKQLVLEASLITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSNQSLFNVTNVAQLGAGQLPAVVAYEAFYQHSFAEQSYDCSRVRVPA	180
Db	121	VNLHCHNTCSNQSLFNVTNVAQLGAGQLPAVVAYEAFYQHSFAEQSYDCSRVRVPA	180

Qy	181	ATLAVTMCVGVGSAALCNAQRWLNFGQDTGNGLAFLDITFHLLERQAVGSGIQPLNEGV	240
Db	181	ATLAVTMCVGVGSAALCNAQRWLNFGQDTGNGLAFLDITFHLLERQAVGSGIQPLNEGV	240
Qy	241	ARCNESQGGDDVATCSCQDCAASCPAARPOALDSTFYLGQMPGSLVLIILCSVFVAVTI	300
Db	241	ARCNESQGGDDVATCSCQDCAASCPAARPOALDSTFYLGQMPGSLVLIILCSVFVAVTI	300
Qy	301	LLVGFVAPARDKSKMVPKGTSLSDKLSFSFTHTLTGQFQCGWGTWASWPLTILVLSV	360
Db	301	LLVGFVAPARDKSKMVPKGTSLSDKLSFSFTHTLTGQFQCGWGTWASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTETLTPVELWSAPNSQARSEKAFHQHFGFPFFRTNQVILTAPNRSSY	420
Db	361	IPVVALAAGLVFTETLTPVELWSAPNSQARSEKAFHQHFGFPFFRTNQVILTAPNRSSY	420
Qy	421	RYDSLLLGPKNFGSILDLLELLELQBLRHLQVWSPQAQRNISILODICYAPLNPNT	480
Db	421	RYDSLLLGPKNFGSILDLLELLELQBLRHLQVWSPQAQRNISILODICYAPLNPNT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDHFLYCANAPLTFKDGITALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDHFLYCANAPLTFKDGITALAL	540
Qy	541	SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLMNYPAGDPRPRLAQAALWBEAFLEEM	600
Db	541	SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLMNYPAGDPRPRLAQAALWBEAFLEEM	600
Qy	601	RAFQRMAGMFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSVSSSRV	660
Db	601	RAFQRMAGMFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSVSSSRV	660
Qy	661	MVDSKATILGGVAVVLGAVMAAGFFSYLIGRSSVLQVVPFLVLSVGDADNIFIVLE	720
Db	661	MVDSKATILGGVAVVLGAVMAAGFFSYLIGRSSVLQVVPFLVLSVGDADNIFIVLE	720
Qy	721	YORLPRPGEPREVHIGRALGRVAPSMLLCSLSBAICFFLGALTPMPAPVTFALTSLGLAV	780
Db	721	YORLPRPGEPREVHIGRALGRVAPSMLLCSLSBAICFFLGALTPMPAPVTFALTSLGLAV	780
Qy	781	ILDFLQMSAFVALLSDSKRQEARSLDVCCCVKQBELPPGQGEGLLGFQKAYAPFL	840
Db	781	ILDFLQMSAFVALLSDSKRQEARSLDVCCCVKQBELPPGQGEGLLGFQKAYAPFL	840
Qy	841	LHWITRGVVLTLFALFGVSLYSMCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVVLTLFALFGVSLYSMCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGAPV	900
Qy	901	YFVTLTGYNFSSEAGMNAICSSAGCNPSPFTQKIQYATEPPEQSYLAIPASSWDDFDIW	960
Db	901	YFVTLTGYNFSSEAGMNAICSSAGCNPSPFTQKIQYATEPPEQSYLAIPASSWDDFDIW	960
Qy	961	LTPSSCCRLYISGPNKDFCSTVNSLNCNKSITWGSVRPSVEQPHKYLFWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDFCSTVNSLNCNKSITWGSVRPSVEQPHKYLFWFLNDRP	1020
Qy	1021	NTKCPKGLAAVSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAANITADL	1080
Db	1021	NTKCPKGLAAVSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAANITADL	1080
Qy	1081	RKVPGTDPAFEPFPTTITNVFYEQYLITLPEGLFMLSICLVPTFAVSCLLGLDLSGLL	1140
Db	1081	RKVPGTDPAFEPFPTTITNVFYEQYLITLPEGLFMLSICLVPTFAVSCLLGLDLSGLL	1140
Qy	1141	NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAIKPTWLER	1200
Db	1141	NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAIKPTWLER	1200
Qy	1201	AKEATTISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLGLHLGLVFLPV	1260
Db	1201	AKEATTISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLGLHLGLVFLPV	1260
Qy	1261	ILSYVGPVNPALALEQKRAEEAAVAVVASCNHPHSRVSTADNIYVNHSPFEGSIKAGA	1320

Db 1261 ILSYGPDPNPALEQKAEAAVAVMVASCPNHPRSVSTADNIVVNSPFGSKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
RESULT 6
AAM79168
ID AAM79168 standard; protein; 1359 AA.
XX
AC AAM79168;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1830.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR N-PSDB; AAK52301.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 20; Page 4213-4216; 5221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKOLVLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKOLVLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSIFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSIFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTMCVGYGSALCNAQRWLNFGQDGTGNGLAPLIDITFHLLERPCQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCVGYGSALCNAQRWLNFGQDGTGNGLAPLIDITFHLLERPCQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVATCQDCQCAASCPAIAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCQDCQCAASCPAIAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWVDPKKGTSLSLSDKLSFSTHTLLGQFFQSGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWVDPKKGTSLSLSDKLSFSTHTLLGQFFQSGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHQHFGPFFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHQHFGPFFRTNQVILTAPNRSY 420
QY 421 RYDSLILGPKNFGSILDLLELELQERLHLQWSPQAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSLILGPKNFGSILDLLELELQERLHLQWSPQAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFONNETLLLTANOTLMGQTSQVDKDHLYCANAPLTFCQGTALAL 540
Db 481 SLYDCCINSLLQYFONNETLLLTANOTLMGQTSQVDKDHLYCANAPLTFCQGTALAL 540
QY 541 SCMADYGAPVFPFLAIGYKGYSEAEALMTFSLNNYPAGDPRLAQAKLWEAFLEBM 600
Db 541 SCMADYGAPVFPFLAIGYKGYSEAEALMTFSLNNYPAGDPRLAQAKLWEAFLEBM 600
QY 601 RAPQRMAGMFQVTPAERSLEDEINRTABDLPIFATSYIVIFLYISLALSGYSWSRV 660
Db 601 RAPQRMAGMFQVTPAERSLEDEINRTABDLPIFATSYIVIFLYISLALSGYSWSRV 660
QY 661 MVDKATILGCGVAVVLGAVMAAGFPFSLVIGTRSSLVILQVVPFLVLSGADNIFVLE 720
Db 661 MVDKATILGCGVAVVLGAVMAAGFPFSLVIGTRSSLVILQVVPFLVLSGADNIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGLALTMPAVRTPALTSGLA 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGLALTMPAVRTPALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVKPQELPPPGQGGELLLGFPQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVKPQELPPPGQGGELLLGFPQKAYAPPL 840
QY 841 LHWITRGVVALFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVALFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKQYATEPPEQSYLAIPASSWDDDFIDW 960
Db 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKQYATEPPEQSYLAIPASSWDDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDQVL-----ASRFMAYH 1053

1054 KPLKNSQDYTEALRAARELAANITADLRKVKFGTDPFAFEVPPYTTITNVFYQYLTILPEGL 1113
1081 KPLKNSQDYTEALRAARELAANITADLRKVKFGTDPFAFEVPPYTTITNVFYQYLTILPEGL 1140
1114 FMLSCLVPTFAVSCLLGLDLRGLNLLSIVMLVDVTVGFMAWDISYNAVSLINLVS 1173
1141 FMLSCLVPTFAVSCLLGLDLRGLNLLSIVMLVDVTVGFMAWDISYNAVSLINLVS 1200
1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVANTNLPGLVLGLAKAQ 1233
1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVANTNLPGLVLGLAKAQ 1260
1234 LIQIFFFRLNLLITLLGLLHGLVFLPVLVSYPDPNPALALEQKRAEEAAVAVWVASCP 1293
1261 LIQIFFFRLNLLITLLGLLHGLVFLPVLVSYPDPNPALALEQKRAEEAAVAVWVASCP 1320
1294 NHPSRVSTADNTIYNHSPGSIKGAISNFPNNGRQF 1332
1321 NHPSRVSTADNTIYNHSPGSIKGAISNFPNNGRQF 1359

RESULT 7
ADJ27337 standard; protein; 1359 AA.
XX
AC ADJ27337;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human NPC1L1.
XX
KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SBD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO2004009772-A2.
XX
XX 29-JAN-2004.
XX
PF 17-JUL-2003; 2003WO-US022467.
XX
PR 19-JUL-2002; 2002US-0397442P.
XX
XX (SCHE) SCHERING CORP.
XX
XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX
XX WPI; 2004-132945/13.
XX
XX N-ESDB; ADJ27336.
XX
XX GENBANK; AF192522.
XX
XX New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting
XX or identifying antagonists of NPC1L1 for inhibiting intestinal
XX cholesterol absorption in a subject, or for treating elevated serum
XX cholesterol or stroke.
XX
XX Example 7; SEQ ID NO 44; 125pp; English.
XX
XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1
XX (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains
XX a motif ADJ27331 which acts as a trans-golgi network to plasma membrane
XX transport signal, and which exhibits limited tissue distribution and
XX gastrointestinal abundance. The human NPC1L1 promoter sequence contains a
XX sterol regulated element binding protein 1 (SREBP1) binding consensus
XX sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
XX sensing domain (SBD) which is involved in sensing cholesterol levels,
XX possibly by a mechanism which involves direct cholesterol bonding. NPC1L1

has 42% amino acid sequence homology to human NPC1, a receptor responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or polynucleotides are useful for detecting or identifying antagonists of NPC1L1, which can be used for inhibiting intestinal cholesterol absorption in a subject, or for treating medical conditions including elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary heart disease, stroke or arteriosclerosis.

Query Match 99.5%; Score 6872.5; DB 8; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
SQ Sequence 1359 AA;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFVDECKNPGLSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFVDECKNPGLSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
DB 121 VNLCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITTHLLEPGQAVSGIQPLNEGV 240
DB 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITTHLLEPGQAVSGIQPLNEGV 240
QY 241 ARCHESQDDVATCSCQDCAASCPAIAARPOALDSTFYLGQMPGSLVLIILCSVFAVVTTI 300
DB 241 ARCHESQDDVATCSCQDCAASCPAIAARPOALDSTFYLGQMPGSLVLIILCSVFAVVTTI 300
QY 301 LLVGFRAVAPARDKSMVDPKKGTSLSKLSSTHTILGQFQCGTGWASWPLTILVLSV 360
DB 301 LLVGFRAVAPARDKSMVDPKKGTSLSKLSSTHTILGQFQCGTGWASWPLTILVLSV 360
QY 361 IPVALAAGLVPTETLTDPVELWSAPNSQARSEKAFHQHFGPPFRTNQVILTAPNRSY 420
DB 361 IPVALAAGLVPTETLTDPVELWSAPNSQARSEKAFHQHFGPPFRTNQVILTAPNRSY 420
QY 421 RYDSILLGPKNFGSILDLLELLELQERLRLHQLVWSPEAQRNISLQDICYAPLNPDNT 480
DB 421 RYDSILLGPKNFGSILDLLELLELQERLRLHQLVWSPEAQRNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDVMDKHFLYCANAPLTPKDGTTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDVMDKHFLYCANAPLTPKDGTTALAL 540
QY 541 SCWADYGAPVPPFLAIGYKGYDSEAEALIMTFSINNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCWADYGAPVPPFLAIGYKGYDSEAEALIMTFSINNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRRMAGMFQVTFPFAERSLEDEINRTTAEDLPFPATSYIIVIFLYISALGYSYSSWSRV 660
DB 601 RAFQRRMAGMFQVTFPFAERSLEDEINRTTAEDLPFPATSYIIVIFLYISALGYSYSSWSRV 660
QY 661 MVDKATLGLGVAVVGLGAVMAAGFFSYLGRSSIVILQVVPFLVLSYGADNIFIFVLE 720
DB 661 MVDKATLGLGVAVVGLGAVMAAGFFSYLGRSSIVILQVVPFLVLSYGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREHIGRALGRVAPSMJLCSLSEATCFFLGALTTPMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREHIGRALGRVAPSMJLCSLSEATCFFLGALTTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCKVQPELPPPGQEGLLGFFOKAVAPFL 840
DB 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCKVQPELPPPGQEGLLGFFOKAVAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

QY 901 YFVTTLTGYNFSEAGNNAICSSAGCNFSTQKIQVATEFPQSYLAIPASSWVDDFDW 960
 DB |||||
 QY 901 YFVTTLTGYNFSEAGNNAICSSAGCNFSTQKIQVATEFPQSYLAIPASSWVDDFDW 960
 DB |||||
 QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCKNCMSITMGSVRPSVEQPHKYLIPWFNDRP 1020
 DB |||||
 QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCKNCMSITMGSVRPSVEQPHKYLIPWFNDRP 1020
 DB |||||
 QY 1021 NIKPKGGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
 DB |||||
 QY 1021 NIKPKGGLAAYSTSVNLTSDQVLTVAIILSPRLSEYGTISAHCNVLLDSASFPMAYH 1080
 DB |||||
 QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPVPTITNVFYEQYLTILPEGL 1113
 DB |||||
 QY 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPVPTITNVFYEQYLTILPEGL 1140
 DB |||||
 QY 1114 FMLSICLVPTFAVSCLLGLDLSRLSGLANLLSVMTILVDTVGPALWDISYNAVSILNLS 1173
 DB |||||
 QY 1141 FMLSICLVPTFAVSCLLGLDLSRLSGLANLLSVMTILVDTVGPALWGISYNAVSILNLS 1200
 DB |||||
 QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233
 DB |||||
 QY 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260
 DB |||||
 QY 1234 LIQIPFFRLNLLITLLGLLGLVFLPVILSVYVGPDPNPALALEQKRAEEAAVAVMSACP 1293
 DB |||||
 QY 1261 LIQIPFFRLNLLITLLGLLGLVFLPVILSVYVGPDPNPALALEQKRAEEAAVAVMSACP 1320
 DB |||||
 QY 1294 NHPSRVSTADNLYNHSPEGSTKGAISNLFNPNNGRQF 1332
 DB |||||
 QY 1321 NHPSRVSTADNLYNHSPEGSTKGAISNLFNPNNGRQF 1359
 DB |||||

RESULT 8

ADY60780

ID ADY60780 standard; protein; 1359 AA.

XX AC

XX ADY60780;

XX DT

XX 19-MAY-2005 (first entry)

XX DE

XX Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 44.

XX KW

Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
 Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
 metabolic disorder; atherosclerosis; cardiovascular disease;
 coronary artery disease; cerebrovascular ischemia; neurological disease;
 arteriosclerosis; transgenic animal.

XX OS

XX Homo sapiens.

XX PN

XX WO2005015988-A1.

XX XX

XX 24-FEB-2005.

XX PD

XX 16-DEC-2003; 2003WO-US040113.

XX PF

XX 17-JUL-2003; 2003US-00621758.

XX PR

XX 22-AUG-2003; 2003US-00646301.

XX PR

XX 16-SEP-2003; 2003US-00663208.

XX XX

XX (SCHE) SCHERING CORP.

XX XX

XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX PI

XX WPI; 2005-284403/29.

XX DR

XX N-PSDB; ADY60779.

XX XX

XX Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of

XX PT

XX NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for

XX PT

XX treating hyperlipidemia or atherosclerosis in mouse, rat or human.

XX XX

PS Disclosure; SEQ ID NO 44; 146pp; English.

XX The invention relates to an isolated Niemann-Pick C1-like protein 1
 CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
 CC amino acid sequence appearing as ADY60738,ADY60740 or ADY60748 (rat,
 CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
 CC encoding NPC1L1 (appearing as ADY60737,ADY60739 or ADY60747), a
 CC recombinant vector comprising the polynucleotide, a host cell comprising
 CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
 CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
 CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
 CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
 CC produce any functional NPC1L1 protein), an offspring or progeny of the
 CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
 CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
 CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject,
 CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
 CC dosage form, and information indicating that NPC1L1 is a target of
 CC ezetimibe or substituted azetidinone), decreasing the level of intestinal
 CC sterol or Salpha-sterol absorption in a subject (involving reducing the
 CC level of expression of NPC1L1 in the subject), identifying an antagonist
 CC of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a
 CC functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist
 CC of NPC1L1 which is useful for inhibiting or decreasing the level of
 CC NPC1L1 mediated sterol or Salpha-sterol uptake in a subject such as
 CC mouse, rat or human. The level of expression of NPC1L1 in the subject is
 CC reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is
 CC useful for screening a sample for intestinal sterol or Salpha-sterol
 CC absorption antagonist. The NPC1L1 antagonists identified are useful for
 CC treating medical conditions caused or mediated by NPC1L1, e.g.,
 CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
 CC arteriosclerosis. The present sequence represents a mammalian NPC1L1
 CC protein.

XX SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 9; Length 1359;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLLLRLAQSEPTTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
 DB |||||
 QY 1 MAEAGLRGWLWALLLLRLAQSEPTTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
 DB |||||
 QY 61 TPARKITGDHLILLOKICPRLYTCGNTQACCSAKOLVSLEASLSITKALLTRCPACSNDF 120
 DB |||||
 QY 61 TPARKITGDHLILLOKICPRLYTCGNTQACCSAKOLVSLEASLSITKALLTRCPACSNDF 120
 DB |||||
 QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
 DB |||||
 QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
 DB |||||
 QY 181 ATLAVGTWCGYVGSALCNQARWLNFGQDTGNGLAPLDITPHLLRPGQAVGSGIQPLNEGV 240
 DB |||||
 QY 181 ATLAVGTWCGYVGSALCNQARWLNFGQDTGNGLAPLDITPHLLRPGQAVGSGIQPLNEGV 240
 DB |||||
 QY 241 ARCNEQGDVVATCSQDCAACSCPAIARPOALDSTFYLGQMPGSLVLIIILCSVPFVVTTI 300
 DB |||||
 QY 241 ARCNEQGDVVATCSQDCAACSCPAIARPOALDSTFYLGQMPGSLVLIIILCSVPFVVTTI 300
 DB |||||
 QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPFQGTWGTWASWPLTILVLSV 360
 DB |||||
 QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPFQGTWGTWASWPLTILVLSV 360
 DB |||||
 QY 361 IPVVALAAGLVFTTETDTPVELWSAPNSQARSEKAFHDQHPGPPFRNQVLTAPNRSY 420
 DB |||||
 QY 361 IPVVALAAGLVFTTETDTPVELWSAPNSQARSEKAFHDQHPGPPFRNQVLTAPNRSY 420
 DB |||||
 QY 421 RYDSLLGPKNFSGILDLDLLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
 DB |||||
 QY 421 RYDSLLGPKNFSGILDLDLLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
 DB |||||
 QY 481 SLYDCCINSLLIQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDTALAL 540
 DB |||||

Db 481 SLIDCCNSLLQYFQNNRTLLLTANQTLMGTSQVDKDHFLYCANAPLTFKDGATALL 540
Qy 541 SCMDYGAPEPPFPIAIGYKGYKDYSEBALIMTFESINNYAGDPRLAQAKLWEAFLEEM 600
Db 541 SCMDYGAPEPPFPIAIGYKGYKDYSEBALIMTFESINNYAGDPRLAQAKLWEAFLEEM 600
Qy 601 RAFORWAGMFOVTFMAERSLEDEINRTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660
Db 601 RAFORWAGMFOVTFMAERSLEDEINRTAEDLPIFATSYIVIFLYISLALGSYSWSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVPPFLVSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVPPFLVSVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQEARSLDVCCCKVQELPPPGQGBGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEARSLDVCCCKVQELPPPGQGBGLLGFQKAYAPFL 840
Qy 841 LHMITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGPV 900
Db 841 LHMITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGPV 900
Qy 901 YFVTLGYNFSSEAGNAICSSAGCNFSFTOKIOYATEFPQSYLAIPASSWVDDFDW 960
Db 901 YFVTLGYNFSSEAGNAICSSAGCNFSFTOKIOYATEFPQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Qy 1021 NIKCPKGLAAVSTSVNLTSDDGVL-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAVSTSVNLTSDDGVL-----ASRPMAYH 1053
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNVFYEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNVFYEQVLTILPEGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWLGISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWLGISYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKENTISMGSAVPAGVAMTNLPGLVGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKENTISMGSAVPAGVAMTNLPGLVGLAKAQ 1260
Qy 1234 LIQIFPRMLNLLITLGLLHGLVFLPVILSYVGFDPVNPALALQKRAEAAVAVMVASCP 1293
Db 1261 LIQIFPRMLNLLITLGLLHGLVFLPVILSYVGFDPVNPALALQKRAEAAVAVMVASCP 1320
Qy 1294 NHPSRVSTADNIYVNHFSFEGSIKAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHFSFEGSIKAGAI SNFLPNNGRQF 1359

RESULT 9
AEB93610
ID AEB93610 standard; protein; 1359 AA.
XX
AC AEB93610;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human NPC1L1 protein.
XX
KW Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;
KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
KW arteriosclerosis; db.

XX OS Homo sapiens.
XX WO2005069900-A2.
XX
XX 04-AUG-2005.
XX
XX 14-JAN-2005; 2005WO-US001469.
XX PF
XX 16-JAN-2004; 2004US-0537341P.
XX PR
XX (MERI) MERCK & CO INC.
XX
XX Garcia-Calvo M;
XX
XX WPI; 2005-564070/57.
XX N-PSDB; AEB93609.
XX GENBANK; AF192522.
XX
XX Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by
XX contacting NPC1L1 with detectably labeled substituted 2-azetidinone
XX glucuronide and a candidate compound and determining if compound binds to
XX human NPC1L1.
XX
XX Disclosure; SEQ ID NO 44; 215pp; English.
XX
XX The invention relates to identifying a ligand of NPC1L1. The method
XX involves contacting human NPC1L1 with a detectably labeled substituted 2-
XX azetidinone glucuronide and a candidate compound and determining whether
XX the candidate compound binds to human NPC1L1. In identifying a ligand of
XX NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10
XX nM or lower. The detectably labeled substituted 2-azetidinone glucuronide
XX is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands
XX are useful for stimulating or blocking the activity of NPC1L1, and for
XX treating conditions caused or mediated by NPC1L1. It is useful for
XX reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart
XX disease, stroke, or arteriosclerosis. The present sequence represents the
XX amino acid sequence of human NPC1L1 as disclosed under GenBank accession
XX number AF192522.
XX
XX Sequence 1359 AA;
XX
XX Query Match 99.5%; Score 6872.5; DB 9; Length 1359;
XX Best Local Similarity 97.9%; Pred. No. 0;
XX Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
XX
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGVCFAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGVCFAFYDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLILLLOKI CPERLYTGNTQACCSAKQLVSLSEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLLOKI CPERLYTGNTQACCSAKQLVSLSEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLVGTMCVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLPEQQAQVSGIQPLNEGV 240
Db 181 ATLVGTMCVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLPEQQAQVSGIQPLNEGV 240
Qy 241 ARCNEQGGDDVATCSQDCAASCAPAIARPAQALDSTFYLQMPGSLVLIILCSVFVAVTI 300
Db 241 ARCNEQGGDDVATCSQDCAASCAPAIARPAQALDSTFYLQMPGSLVLIILCSVFVAVTI 300
Qy 301 LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTLTTDVPVELWSAPNSQARSEKAFHDQHFQFFFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTLTTDVPVELWSAPNSQARSEKAFHDQHFQFFFRTNQVILTAPNRSSY 420


```
QY 421 RYDSSLLGPKNPSGILDLDLLELELEQLERLRLHQLVMSPEAQRNISLQDICVAPLNPDNT 480
Db 421 RYDSSLLGPKNPSGILDLDLLELELEQLERLRLHQLVMSPEAQRNISLQDICVAPLNPDNT 480
QY 481 SYDDCCINSLLQYFQNNRTLLATNTQTLTGOTSDQVMDKDFLYCANAPLTKDGTALAL 540
Db 481 SYDDCCINSLLQYFQNNRTLLATNTQTLTGOTSDQVMDKDFLYCANAPLTKDGTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
QY 601 RAFORMAGMFWQVTFPAERSLSEDEINRTTAEDLPFATSYIYIPLYISLALSYSSWSRV 660
Db 601 RAFORMAGMFWQVTFPAERSLSEDEINRTTAEDLPFATSYIYIPLYISLALSYSSWSRV 660
QY 661 MYDSKATLGLGCVAVVLGAVMAAMGFFSYLGRSSILVLOVVPFLVLSVGADNIFVLE 720
Db 661 MYDSKATLGLGCVAVVLGAVMAAMGFFSYLGRSSILVLOVVPFLVLSVGADNIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAPVTRTALTSGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAPVTRTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKPQELPPPPQCGEGLLGFFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKPQELPPPPQCGEGLLGFFQKAYAPPL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YPVTTILGNFSSEAGNNAICSSAGCNFSFTOKIYATFEPQSYLAIPASSWVDVDFIDW 960
Db 901 YPVTTILGNFSSEAGNNAICSSAGCNFSFTOKIYATFEPQSYLAIPASSWVDVDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITWGSVRPSVEQFHKLPLWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITWGSVRPSVEQFHKLPLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSNLTSDQVLT-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAYSTSNLTSDQVLT-----ASRPMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTIITNVFEQVLTILPGL 1113
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTIITNVFEQVLTILPGL 1113
QY 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTIITNVFEQVLTILPGL 1140
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTIITNVFEQVLTILPGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLLMLLSYVMTLVDTVGPMLWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRLSGLLMLLSYVMTLVDTVGPMLWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260
QY 1234 LIQIFEFRLNLTILGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEAAVAVMVASCP 1293
Db 1261 LIQIFEFRLNLTILGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEAAVAVMVASCP 1320
QY 1294 NHPSRVSTADNTIYVNHSEFSGIKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNTIYVNHSEFSGIKGAGAI SNFLPNNGRQF 1359
```

RESULT 10

ABG22693

ID ABG22693 standard; protein; 1344 AA.

XX

AC ABG22693;

XX

DT 18-FEB-2002 (first entry)

```
XX Novel human diagnostic protein #22684.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
OS Homo sapiens.
PN WO200175067-A2.
PD 11-OCT-2001.
PF 30-MAR-2001; 2001WO-US008631.
PR 31-MAR-2000; 2000US-00540217.
PX 23-AUG-2000; 2000US-00649167.
PA (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS86880.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 53052; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic
amino acid sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1344 AA;
SQ
Query Match 94.6%; Score 6536; DB 4; Length 1344;
Best Local Similarity 96.7%; Pred No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;
QY 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAPFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAPFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLTYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLTYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSPNQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLCHNTCSPNQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAVGTWCGYVGSALCNQAORWLNFGQDGTGNGLAPLDTTFHLLFPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTWCGYVGSALCNQAORWLNFGQDGTGNGLAPLDTTFHLLFPGQAVGSGIQPLNEGV 240
```

Db 181 ATLAVGTCVGYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLPEGQAVGSGIOPLNEGV 240

Qy 241 ARCHESQDDVATCSCQDCAASCP--AIARPOALDSTFYLGQMPGSLVLIILICSVFVAV 298

Db 241 ARCHESQDDVATCSCQDCAASCPXPAPRSTPPSTW--RQMPGSLVLIILICSVFVAV 298

Qy 299 TILLVGFVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASPLTILVL 358

Db 299 TILLVGFVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASPLTILVL 358

Qy 359 SVIPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQFPGPFRTNQVILTAPNRS 418

Db 359 SVIPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQFPGPFRTNQVILTAPNRS 418

Qy 419 SYRYSLLGPKNFGSGILDLDLLELLELQERLRLHQLWSPQAQNSLQDICYAPLNPD 478

Db 419 SYRYSLLGPKNFGSGILDLDLLELLELQERLRLHQLWSPQAQNSLQDICYAPLNPD 478

Qy 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGTL 538

Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGTL 538

Qy 539 ALSCHADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 598

Db 539 ALSCHADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 598

Qy 599 EMRAFQRRMAGMFQVFTTAERSLEDEINRTTAEDLPFATSIVIVIFLYISLALGYSYWS 658

Db 599 EMRAFQRRMAGMFQVFTTAERSLEDEINRTTAEDLPFATSIVIVIFLYISLALGYSYWS 658

Qy 659 RVWDSKATLGLGGVAVTVGAWAAGPESYLGIRSSVLQVPLVLSVGADNIFIV 718

Db 659 RVWDSKATLGLGGVAVTVGAWAAGPESYLGIRSSVLQVPLVLSVGADNIFIV 718

Qy 719 LEYQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSL 778

Db 719 LEYQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSL 778

Qy 779 AVILDFLLQWAFVALLSDSKRQEARLDVCCVQBELPPGQEGILLGFFQKAYAP 838

Db 779 AVILDFLLQWAFVALLSDSKRQEARLDVCCVQBELPPGQEGILLGFFQKAYAP 838

Qy 839 FLHLWITRGVV-----LLFLALFGVSLYSNMCHTSVGLDQELALPKDSYLLDYFLFNRYF 894

Db 839 FLHLWITRGVVPSQLLFLALFGVSLYSNMCHTSVGLDQELALPKDSYLLDYFLFNRYF 898

Qy 895 EVGAPVYFVTTLGYNFSSEAGMNAICSSAGCANNFSFTQKIQVATEPPEQSYLAIPASSWV 954

Db 895 EVGAPVYFVTTLGYNFSSEAGMNAICSSAGCANNFSFTQKIQVATEPPEQSYLSLPASSWV 958

Qy 955 DDFIDMLTPSSCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRPSVQFHKYLPW 1014

Db 955 DDFIDMLTPSSCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRPSVQFHKYLPW 1018

Qy 1015 FLNDRPNKCPKGLAAYSTVNLTSQGVLASRFMAYHKPLKNQSDYTEALRAARELAA 1074

Db 1015 FLNDRPNKCPKGLAAYSTVNLTSQGVLASRFMAYHKPLKNQSDYTEALRAARELAA 1078

Qy 1075 NITADLRKVPGTDPAPFVFPYTTITNVFYEQLYTLPEGLPMLSLCLVPTFAVSCLLGLD 1134

Db 1075 NITADLRKVPGTDPAPFVFPYTTITNVFYEQLYTLPEGLPMLSLCLVPTFAVSCLLGLD 1138

Qy 1135 LRSGLLNLLSIVMLVDTVGFMALWDISYNVNSLINLVSAGMVSFVSHITSFALSTK 1194

Db 1135 LRSGLLNLLSIVMLVDTVGFMALWDISYNVNSLINLVSAGMVSFVSHITSFALSTK 1198

Qy 1195 PTWLERAKEATISMGSAVFAGVAMTNLPGLIIVLGLAKAQIQLIFFFLNLIITLGLLHG 1254

Db 1195 PTWLERAKEATISMGSAVFAGVAMTNLPGLIIVLGLAKAQIQLIFFFLNLIITLGLLHG 1258

Qy 1255 LVFLPVLISYVGPDPNPALEQKRABEA-----VAAMVWASCPNHSRVST 1301

Db 1255 LVFLPVLISYVGPDPNPALEQKRABEGGGSHGFLAQITPPESPQLTTSMT 1313

RESULT 11

ADJ27295

ID ADJ27295 standard; protein; 1331 AA.

XX AC ADJ27295;

XX AC ADJ27295;

DT 20-MAY-2004 (first entry)

XX DE Rat NPC1L1.

XX NIemann-Pick disease; type CI; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promotor;

XX sterol regulated element binding protein 1; SREBP1;

XX binding consensus sequence; transmembrane domain; sterol-sensing domain; SSD; cholesterol; NPC1; receptor; Niemann-Pick CI disease; intestinal; cholesterol absorption; serum cholesterol; hyperlipidaemia; atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX Rattus sp.

XX WO2004009772-A2.

XX 29-JAN-2004.

XX 17-JUL-2003; 2003WO-US022467.

XX 19-JUL-2002; 2002US-0397442P.

XX (SCHE) SCHERING CORP.

XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

DR WPI; 2004-132945/13.

DR N-PSDB; ADJ27294.

XX New Niemann-Pick disease, type CI, gene-like 1 polypeptide, for detecting or identifying antagonists of NPC1L1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

XX Claim 1; SEQ ID NO 2; 125pp; English.

XX This sequence represents a Niemann-Pick disease, type CI, gene-like 1 (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and gastrointestinal abundance. The human NPC1L1 promoter sequence contains a sterol regulated element binding protein 1 (SREBP1) binding consensus sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-sensing domain (SSD) which is involved in sensing cholesterol levels, possibly by a mechanism which involves direct cholesterol binding. NPC1L1 has 42% amino acid sequence homology to human NPC1, a receptor responsible for Niemann-Pick CI disease. The NPC1L1 polypeptides or polynucleotides are useful for detecting or identifying antagonists of NPC1L1, which can be used for inhibiting intestinal cholesterol absorption in a subject, or for treating medical conditions including elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary heart disease, stroke or arteriosclerosis.

XX Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 8; Length 1331;

Best Local Similarity 77.8%; Pred. No. 0;

Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAEAGLRGMLLWALLRLAQSEPYTIHQPGYCAFDCEGKPELSGSLMTLSNVSLCN 60

Db 1 MAAAWL-GWLLWALLLSAAQGLYTPKHEAGVCTFYEECGKPELSGSLMTLSNVSLCN 59

Qy 61 TPARKITGPHLLIILQKICPRLVTPGNQ-ACCSAKOLVSLASLSITKALLTRCPACSDN 119

Db 60 TPAHVTGEHLALQRICPLRYNGPNTTFACCKTKQLLSLESSMSITKALLTRCPACSDN 119
QY 120 FVNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVRVPA 179
Db 120 FVSLHCHNTCSNPQSLFINVTRVVERGAGPPAVVAEYAFYQHSFAEQSYDSCSRVRIPA 179
QY 180 AATLAVGTCGVVGSALCNAQRMWLNFGDGTGNGLAFLDITFHLLEPGQAVGSGIQPLNEG 239
Db 180 AASLAVGSMGVVGSALCNAQRMWLNFGDGTGNGLAFLDITFHLLEPGQALPDGIQPLNGK 239
QY 240 VARNESQGDVATCSCODCAAPALARPQALDSTFYLGOMPGSLVLIILCSVFAVVT 299
Db 240 IAPCNESQGDSDVSCVSCODCAAPCVIPPPPEARLPSPFTMGFMFWLALIIITFAVFLVS 299
QY 300 ILVGFVRVAPARDKSKMVPDKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLS 359
Db 300 AVLVLRLVSVNRNKNKAEGQAPKLPKHKLSPHILGRFFQNMGTFRVASWPLTVLALS 359
QY 360 VIPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDOHFQPPFRTRNQVILTAPNRSS 419
Db 360 FIVVIALAAGLTPIELTDPVELWSAPNSQARSEKSFHDEHFGFFFRTRNQIFVTAARNSS 419
QY 420 YRYDSLILGPKFNGSILDLLELLELQERLRLHLOWSPQAORNISLODICVAPLNPDN 479
Db 420 YKYSLLSAGSKNFSGLSLDLELLELQERLRLHLOWSPQAORNISLODICVAPLNPDN 479
QY 480 TSLYDCCINSLLQYFQNNRLLLTANQTLMGQTSOVDKMDHFLYCANAPLTPKQGTALA 539
Db 480 TSLSDCCVNSLLQYFQNNRLLLTANQTLMGQTSOVDKMDHFLYCANAPLTPKQGTSLA 539
QY 540 LSCMADYGAPVPFPLAIGYKGDYSEAEALIMTFSNNYPAGDPRLAQAUKWEAFLE 599
Db 540 LSCMADYGAPVPFPLAIGYKGDYSEAEALIMTFSNNYPAGDPRLAQAUKWEAFLE 599
QY 600 MRAFORRMAGMFOVTFATRSLEDEINRTTAEPLIPATSYIVIVLYISLALGSSYSWR 659
Db 600 MESFORNTSKFOVAFSAKRSLEDEINRTTQDLPVFAVSYIIVLYISLALGSSYSRCSR 659
QY 660 VMVDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFVL 719
Db 660 VAVESKATILGCGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFVL 719
QY 720 EYQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSLGA 779
Db 720 EYQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSLGA 779
QY 780 VILDFFLQMSAFVALLSLDSKQEAERLDVCCCKVQBELPPQCGGLILGFPKAYAPF 839
Db 780 IILDFFLQMTAFVALLSLDSKQEAERLDVCCCKVQBELPPQCGGLILGFPKAYAPF 839
QY 840 LLHWITRGVVLILFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPRVGPAP 899
Db 840 LLHRFTIRPVVMLLFLTLFGANLYMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGPP 899
QY 900 VYFVTTLGNFSEAGMNAICSSAGCNFSTFKIQIYATEPPEQSYLAIPASSWDDFD 959
Db 900 VYFVTTSGFNFSSEAGMNAICSSAGCNFSTFKIQIYATEPPEQSYLAIPASSWDDFD 959
QY 960 WLTP-SSCCRLYISGNKDFCEPSTVNSLNCNKMCHITWGSVRPSVEGPHKYLFWFLND 1018
Db 960 WLTPSSCCRLYIRGPHKDFCEPSTVNSLNCNKMCHITWGSVRPSVEGPHKYLFWFLND 1019
QY 1019 RPNKCPKGLAAYSTSVNLTSQVLSRFRMAYHKPLKNSODYTBALRAARELANITA 1078
Db 1020 PNIRCPKGLAAYSTSVNLSDGQVIAQFMAHKPLKNSQDFTBALRASULLANIITA 1079
QY 1079 DLKRVGTPDAFEVFPYITINVFYEQYLITLPEGLPMLSLCLVPTFAVSCLLGLDLRG 1138
Db 1080 DLKRVGTPDNFVFPYITINVFYEQYLITLPEGLPMLSLCLVPTFAVSCLLGLDMCSG 1139
QY 1139 ILNLSIVMLVDVTGFMALWISYNAVSLINLVSAGVMSVFPVSHITRSPASTKPTWL 1198
Db 1140 ILNLSIIMILVDITGLMAVWGISYNAVSLINLVTAVGMSVFPVSHITRSPASTKPTRL 1199

QY 1199 ERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLTIQIFFRNLNLIITLLGLHGLVFL 1258
Db 1200 ERAKDATVFMGSAVFAGVAMTNLPGILVLGLAKAQLTIQIFFRNLNLIITLLGLHGLVFL 1259
QY 1259 PVLSTVGDVDPNALLBOKRABEAAVAVWVASCNHPHSRVSTADNINYNHSESGSI-KG 1317
Db 1260 PVLSTVGDVDPNALLBOKRABEAAVAVWVASCNHPHSRVSTADNINYNHSESGSI-KG 1316
QY 1318 AGAISNPLNNGROF 1332
Db 1317 ANAARSLPKSDORF 1331
RESULT 12
ADY60738
ID ADY60738 standard; protein; 1331 AA.
XX ADY60738;
AC
XX
XX 19-MAY-2005 (first entry)
XX Rat Niemann-Pick C1-like protein 1 (NPC1L1) protein.
DE
XX Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX
OS Rattus sp.
XX
XX WO2005015988-A1.
XX
XX 24-FEB-2005.
XX
XX 16-DEC-2003; 2003WO-US040113.
XX
XX 17-JUL-2003; 2003US-00621758.
XX 22-AUG-2003; 2003US-00646301.
XX 16-SEP-2003; 2003US-00663208.
XX
XX (SCHE) SCHERING CORP.
XX
XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX
XX WPI; 2005-284403/29.
XX N-PSDB; ADY60737, ADY60746.
XX
XX Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
XX
XX Claim 30; SEQ ID NO 2; 146pp; English.
XX
XX The invention relates to an isolated Niemann-Pick C1-like protein 1
CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,
CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
CC recombinant vector comprising the polynucleotide, a host cell comprising
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
CC produce any functional NPC1L1 protein), an offspring or progeny of the
CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
CC kit (comprising ezetimibe or substituted azeitidine in a pharmaceutical
CC dosage form, and information indicating that NPC1L1 is a target of
CC ezetimibe or substituted azeitidine), decreasing the level of intestinal
CC sterol or Salpha-sterol absorption in a subject (involving reducing the

CC level of expression of NPC111 in the subject), identifying an antagonist
 CC of NPC111 and an isolated mammalian cell which lacks a gene encoding a
 CC functional NPC111 protein. NPC111 is useful for identifying an antagonist
 CC of NPC111 which is useful for inhibiting or decreasing the level of
 CC NPC111 mediated sterol or Salpha-sterol uptake in a subject such as
 CC mouse, rat or human. The level of expression of NPC111 in the subject is
 CC reduced by mutating NPC111 in the subject. The NPC111 knockout mouse is
 CC useful for screening a sample for intestinal sterol or Salpha-sterol
 CC absorption antagonist. The NPC111 antagonists identified are useful for
 CC treating medical conditions caused or mediated by NPC111, e.g.,
 CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
 CC arteriosclerosis. The present sequence represents a mammalian NPC111
 CC protein.
 XX
 SQ Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 9; Length 1331;
 Best Local Similarity 77.8%; Pred. No. 0;
 Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKNPELSGSLTSLNSVCLSN 60
 DB 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKNPELSGSLTSLNSVCLSN 59
 QY 61 TPARKITGDHLLIQICPRLYTGNTO-ACCSAKOLVSLASLSITKALLTRCPACSDN 119
 DB 60 TPARKITGDHLLIQICPRLYTGNTO-ACCSAKOLVSLASLSITKALLTRCPACSDN 119
 QY 120 FYNLHCHNTCSNQSLFINVTRVAQGLPAAVAYEAFYQHSFAEQSYDCSCSRVRPA 179
 DB 120 FYNLHCHNTCSNQSLFINVTRVAQGLPAAVAYEAFYQHSFAEQSYDCSCSRVRPA 179
 QY 180 AATLAVGTWCYVGSALCNAQRLNPGQDTGNGLAPLDTIFHLBPQAVGSGIOPLNKG 239
 DB 180 AATLAVGTWCYVGSALCNAQRLNPGQDTGNGLAPLDTIFHLBPQAVGSGIOPLNKG 239
 QY 240 VARCHESQDDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILCSVPVVT 299
 DB 240 VARCHESQDDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILCSVPVVT 299
 QY 300 ILLVGRVAPADKSKWDPKGTSLSDKLSFSTHTLQGFQGMGTWVASHPLTILVLS 359
 DB 300 ILLVGRVAPADKSKWDPKGTSLSDKLSFSTHTLQGFQGMGTWVASHPLTILVLS 359
 QY 360 VIPVVALAAGLVTELTTPVELWSPNSQARSEKAFHQHFGPPPTPTQVILTAPNRSS 419
 DB 360 VIPVVALAAGLVTELTTPVELWSPNSQARSEKAFHQHFGPPPTPTQVILTAPNRSS 419
 QY 420 YRVDLSLLGKPFSGILDLDLELELOERLRLQVWSPQARNISLODICVAPLNPDN 479
 DB 420 YRVDLSLLGKPFSGILDLDLELELOERLRLQVWSPQARNISLODICVAPLNPDN 479
 QY 480 TSLYDCINSLLOFYQNNRLLLTANQTLMGQTSQVMDKDFLYCANAPLTFKGTALA 539
 DB 480 TSLYDCINSLLOFYQNNRLLLTANQTLMGQTSQVMDKDFLYCANAPLTFKGTALA 539
 QY 540 LSCMADYGAPVFPFLAIGVKGKDYSEAEALIMTFSINNYPADGPRLAQKLWEEAFLEE 599
 DB 540 LSCMADYGAPVFPFLAIGVKGKDYSEAEALIMTFSINNYPADGPRLAQKLWEEAFLEE 599
 QY 600 MRAFORRMAGMFOVTFABRSLEDEINRTTAEOLPIFATSYIVIFYLSIALGYSYSGWR 659
 DB 600 MRAFORRMAGMFOVTFABRSLEDEINRTTAEOLPIFATSYIVIFYLSIALGYSYSGWR 659
 QY 660 VMYDSKATLGLGGVAVVGLAVMAAMGFFSYLIGRISLVLQVVPFLVLSVGADNIFIVL 719
 DB 660 VMYDSKATLGLGGVAVVGLAVMAAMGFFSYLIGRISLVLQVVPFLVLSVGADNIFIVL 719
 QY 720 EYQRLPRRPOEPRVHIGRALGRVAPSMMLCSLSBAICFFLGLATPMPAVRTALTSGLA 779
 DB 720 EYQRLPRRPOEPRVHIGRALGRVAPSMMLCSLSBAICFFLGLATPMPAVRTALTSGLA 779
 QY 780 VILDFLQMSAFVALLSLDSKQAEASRLDVCCCKVQPELPPQGGEGLLIGFPQKAYAPP 839

DB 780 IILDFLQMSAFVALLSLDSKQAEASRLDVCCCKVQPELPPQGGEGLLIGFPQKAYAPP 839
 QY 840 LLHWITGVVLLLFALFGLSVLYSMCHSISVGLDQELALPKDSYLLDYFLFLNRFEVGPAP 899
 DB 840 LLHWITGVVLLLFALFGLSVLYSMCHSISVGLDQELALPKDSYLLDYFLFLNRFEVGPAP 899
 QY 900 VYFVTTLYGNFSSAGNAICSSAGCNFFSFTQKIQVATPEPBOISYLAIPASSWVDDFID 959
 DB 900 VYFVTTLYGNFSSAGNAICSSAGCNFFSFTQKIQVATPEPBOISYLAIPASSWVDDFID 959
 QY 960 WLTP-SSCCRLYISGPNKDKFCPSSTVNSLNCNKCMSITMGSVRPSVQBFHKYLPWFLND 1018
 DB 960 WLTP-SSCCRLYISGPNKDKFCPSSTVNSLNCNKCMSITMGSVRPSVQBFHKYLPWFLND 1018
 QY 1019 RPNIKCPKGGLAAYSTSVNLSDQVLSRPMAYHKPKNSQDYTEALRAARELAANITA 1078
 DB 1020 RPNIRCPKGGLAAYRTSVNLSSDQVLSRPMAYHKPLRNSQDTEALRASRLLAANITA 1079
 QY 1079 DLKRVGTDPDAFEVFPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSQ 1138
 DB 1080 DLKRVGTDPDAFEVFPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSQ 1139
 QY 1139 LNLLSIVMILVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISYKPTWL 1198
 DB 1140 LNLLSIIMILVDTITGLMAVWGISYNAVSLINLVTAVGMSVEFVSHITRSFAVSKPTRL 1199
 QY 1199 ERAKEATISMGSAVPAGVANTNLPGILVGLAKAQLIQIFPRLNLLITLGLHGLVPL 1258
 DB 1200 ERAKDATVFMGSAVPAGVAMTNFPGILVGLAQLIQIFPRLNLLITLGLHGLVPL 1259
 QY 1259 PVLLSYGDPVNPALALROKRAEAAVAVMVASCPNHPHSRVSTADNIVNHSFEFSI-KG 1317
 DB 1260 PVLLSYGDPVNPALALROKRAEAAVAVMVASCPNHPHSRVSTADNIVNHSFEFSI-KG 1316
 QY 1318 AGAISNLFNNGRQF 1332
 DB 1317 ANAESSLPKSDQKF 1331

RESULT 13
 AEB93568
 ID AEB93568 standard; protein; 1331 AA.
 XX AEB93568;
 AC AEB93568;
 XX 06-OCT-2005 (first entry)
 DT 06-OCT-2005 (first entry)
 XX Rat NPC111 protein.
 DE Protein engineering; NPC111; antidiabetic; antiarteriosclerotic;
 KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
 KW arteriosclerosis; rat.
 XX Rattus sp.
 XX WO2005069900-A2.
 XX 04-AUG-2005.
 XX 14-JAN-2005; 2005WO-US001469.
 XX 16-JAN-2004; 2004US-0537341P.
 XX (MERI) MERCK & CO INC.
 XX Garcia-Calvo M;
 XX WPI: 2005-564070/57.
 XX N-PSDB; AEB93567.
 XX Identifying ligand of NPC111 for stimulating the activity of NPC111, by
 PT contacting NPC111 with detectably labeled substituted 2-azetidinone

glucuronide and a candidate compound and determining if compound binds to human NPC1L1.

Example 2; SEQ ID NO 2; 215pp; English.

The invention relates to identifying a ligand of NPC1L1. The method involves contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining whether the candidate compound binds to human NPC1L1. In identifying a ligand of NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10 nM or lower. The detectably labeled substituted 2-azetidinone glucuronide is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands are useful for stimulating or blocking the activity of NPC1L1, and for treating conditions caused or mediated by NPC1L1. It is useful for reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart disease, stroke, or arteriosclerosis. The present sequence represents a rat NPC1L1, a N-glycosylated protein.

Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 9; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELGGSLMTLSNVCSLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGBLYTPKHEAGVCTFEYECGKNPELGGSLTSLSNVCSLSN 59
QY 61 TPARKITGDHLLLOKICPRLVTGNTQ-ACCSAKOLVSLEASLSITKALLTRCPACSDN 119
DB 60 TPARKVTGEHLALLQRIICPRLVNGNTTFACCSTKQLLSSESSITKALLTRCPACSDN 119
QY 120 FVNLCHNTCSNQSLFNVTRVAQAGQAPVAVYAEAFYQHSFAEQSDYSCSRVRVPA 179
DB 120 FVSLCHNTCPDQSLFNVTRVVERGAGEPAVAVYAEAFYQHSFAEQSDYSCSRVRIPA 179
QY 180 AATLAVGTWCGYGSALCNAQRLNFGQDTGNGLAPLDITPHLLPFGQAVGSGIQLNEG 239
DB 180 AASLAVGSMCGYGSALCNAQRLNFGQDTGNGLAPLDITPHLLPFGQALPDGIQLNGK 239
QY 240 VARCNESQGDVATCSQDCAASCPAIARQALDSTFVYLGQWPSLVLIILCSFVAVT 299
DB 240 IAPCNESQGDOSAVCSQDCAASCPIVPPPEALRPSFYMGKMPGWLIIIFTAVFVLS 299
QY 300 ILLVGFVAPARDKSMVDPKKTSGLSKLSPSTHTLLGQFQGGTGWVWASPLTILVLS 359
DB 300 AVLVELRVVSNKKNKAGSQEAPKLPKHKLSPHTILGRFQNGTRVWASPLTILVLS 359
QY 360 VIPVVALAGLVFTLTDTPVELMSAPNSQARSEKAFHDQHPGPPFRTNQVILTPNRSS 419
DB 360 FIVVIALAAGLTFFIELTTPVELMSAPNSQARSEKAFHDQHPGPPFRTNQVILTPNRSS 419
QY 420 YRYDSLILGPKNFSGILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 479
DB 420 YKYDSLILGPKNFSGILDLLELLELLELLELLELLELLELLELLELLELLELLEL 479
QY 480 TSLYDCCNSLLOYPQNNRTLLLTANOTLMGQTSQVDKDHFLYCANAPLTFKDGITALA 539
DB 480 TSLSDCCVNSLLOYPQNNRTLLLTANOTLMGQTSQVDKDHFLYCANAPLTFKDGITALA 539
QY 540 LSCMADYGAPVPPFLAIGYKGKDYSEAEALITMFLSNYPAGDPRLAQAKLWEBAFLKE 599
DB 540 LSCMADYGAPVPPFLAVGGYQGTDYSEAEALITMFLSNYPAGDPRLAQAKLWEBAFLKE 599
QY 600 MRAQORRNAGHQVTFPAAERSLEDBINRTTAEIDLPIFATSVIIVFLYISLALGSYSWSR 659
DB 600 MESFORNTSDRFQVAFSAERSLEDBINRTTQDLFPVFAVSIVFLYISLALGSYSRCSR 659
QY 660 VVWDSKATLGLGGVAVLVGAVMAAGFPFSYLGIRSSLVILQVPPVLVSGADNIFIFVL 719
DB 660 VAVESKATLGLGGVIVLVGAVLAAGFYSLGVPSLVIIQVPPVLVSGADNIFIFVL 719
QY 720 EYQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSGLA 779

DB 720 EYQRLPRPGEPREAHIGRTLGSAVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
QY 780 VILDPELLQWSAFVALLSLSDSKQEARSLDVCCVCPQELPPPPQEGELILGFQKAYAPF 839
DB 780 IILDPELLQWTAFAVALLSLSDSKQEARSDVLCFSTRKLPKPKKBEGLLLRFRKIYAPF 839
QY 840 LLIHMTITRGVLLLFLLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYPRVGPAP 899
DB 840 LLIHRFIRPVVMLLFLTLFGANLYLMCNINVLGDQELALPKDSYLLDYFLFLNRYLEVGP 899
QY 900 VYFVTTLGYNSSSEAGMAICSSAGCNPFPTQKIQTATEPPEQSYLAIIPASSWDDFD 959
DB 900 VYFVTTSFNFSSSEAGMATCSSAGCKSFSLTKIQYASEFPDQSYVAIAASSWDDFD 959
QY 960 WLTP--SSCCRLYISGPNKPCPSVNSINCLKNCSITMGSVRPSVQFHKYLPWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCLKNCHNRTLGPRVPTAEQFHKYLPWFLND 1019
QY 1019 RPNICPKPGGLAAYSTVNLTSQGLASRFMAYHKPNKNSQDYTEALRAARELAANITA 1078
DB 1020 PPNIRCPKPGGLAAVYTSVNLSSDQGVASQFMAVYHKPLRNSQDFTALRASLLAANITA 1079
QY 1079 DLRKVPGTDPAFEPVPTITNVFYEQYLTITLPEGLFMLSCLVPTFAVSCLLGLDLRSG 1138
DB 1080 DLRKVPGTDPNFEVPPYITISNVFYQYLTLPFEGITLALCFVPTFVVCYLLGLDMCSG 1139
QY 1139 LILNLSIYVMIILVDTVGFMAWDISVNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198
DB 1140 LILNLSIYMIILVDTITGLMAVNGISYNVSLINLTVAGMSVEFVSHITRSPAISTKPTRL 1199
QY 1199 ERAKEATTSMGSAVPAGVAMTNLPGILVLGLAKAQIQLFPPRLNLLITLGLLHGLVPL 1258
DB 1200 ERAKDATVPMGSAVPAGVAMTNFPGLILGLPAQALQIQLFPPRLNLLITLGLLHGLVPL 1259
QY 1259 PVILSYVGPDVNPALALEQKRAEAAVAVMVASCPNHPFSRVSTADNIYVNHSEFSGI-KG 1317
DB 1260 PVILSYLGPDVNQALVQBEKLESEA-AVAPBPCPYQSPADADAN--VNYGFAPELAHG 1316
QY 1318 AGATSNFLPNNGRQF 1332
DB 1317 ANAARSLPKSDQRP 1331
RESULT 14
ADJ27305
ID ADJ27305 standard; protein; 1333 AA.
XX AC ADJ27305;
XX DT 20-MAY-2004 (first entry)
XX DE Mouse NPC1L1.
XX KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain; SSB; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal; cholesterol absorption; serum cholesterol; hyperlipidaemia; atherosclerosis; coronary heart disease; stroke; arteriosclerosis.
XX OS Mus sp.
XX FN WO2004009772-A2.
XX PD 29-JAN-2004.
XX PF 17-JUL-2003; 2003WO-US022467.
XX PR 19-JUL-2002; 2002US-0397442P.
XX SCHE) SCHERING CORP.

[illegible]

Tue Apr 11 10:57:55 2006

Qy	1262	LSYVGPDPNFPALALEQKRAEENVAAMVASCNPHPSPRVSTADNIYNHNSPEGS- IKGAGA	1320
Dδ	1263	LSYLGPDPNQALVLEEKLATEA-AWSEPCQPFFPADANTSQDVYNGFNPFPIEINA	1321
Qy	1321	ISNFLPNNGRQF	1332
Dδ	1322	ASSSLPKSDQKF	1333

Search completed: April 7, 2006, 19:16:33
Job time : 208 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:16:49 ; Search time 52 Seconds
(without alignments)
2464.629 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNNGRQF 1332
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2385	34.5	1278	2 T30188	Niemann-Pick C dis
2	1394	20.2	1055	2 T05663	hypothetical prote
3	1347	19.5	1456	2 T15961	hypothetical prote
4	1329	19.2	1170	2 S52525	probable membrane
5	646.5	9.4	1434	2 T30172	transmembrane prot
6	628.5	9.1	1442	2 T18538	patched protein -
7	600	8.7	915	2 S44797	F09G8.4 protein -
8	598.5	8.7	933	2 T25600	hypothetical prote
9	583.5	8.4	1182	2 T13952	membrane protein p
10	573.5	8.3	1003	2 T26746	hypothetical prote
11	572	8.3	1220	2 T18291	patched protein -
12	539.5	7.8	889	2 T29590	hypothetical prote
13	514	7.4	1405	2 T27969	hypothetical prote
14	473.5	6.9	1015	2 T15830	hypothetical prote
15	464.5	6.7	1299	1 S08119	membrane protein p
16	458.5	6.6	955	2 T21612	hypothetical prote
17	436	6.3	956	2 A89153	protein C24B5.3 [i
18	409.5	5.9	936	2 T26521	hypothetical prote
19	407	5.9	800	2 T26683	hypothetical prote
20	399.5	5.8	881	2 T31739	hypothetical prote
21	381	5.5	840	2 T33217	hypothetical prote
22	379	5.5	714	2 T16126	hypothetical prote
23	367	5.3	690	2 T23399	hypothetical prote
24	367	5.3	877	2 T24097	hypothetical prote
25	349	5.1	845	2 T25657	hypothetical prote
26	347.5	5.0	820	2 T32908	hypothetical prote
27	312	4.5	890	2 T22186	hypothetical prote
28	297.5	4.3	871	2 T28706	hypothetical prote
29	294	4.3	983	2 T21213	hypothetical prote

30	276.5	4.0	633	2 S44795	F09G8.3 protein -
31	259.5	3.8	413	2 S28276	hypothetical prote
32	224	3.2	1276	2 T18526	SREBP cleavage act
33	189.5	2.7	1154	2 T48829	related to SREBP c
34	187.5	2.7	932	2 T28820	hypothetical prote
35	186.5	2.7	1227	2 T20370	hypothetical prote
36	179.5	2.6	787	2 H71453	hypothetical prote
37	173	2.5	969	2 T33156	hypothetical prote
38	170	2.5	746	2 A75018	transport integral
39	154	2.2	823	2 B81282	probable ABC-type
40	149	2.2	1011	2 T07712	multidrug resistan
41	149	2.2	1051	2 AG3455	hypothetical prote
42	148.5	2.1	746	2 H84301	multidrug efflux p
43	146	2.1	1041	2 AC0423	antibiotic transpo
44	143	2.1	724	2 H69780	acriflavin resista
45	141.5	2.0	1049	2 AP0561	

ALIGNMENTS

RESULT 1

T30188
Niemann-Pick C disease protein - mouse
N:Alternate names: NCP1 protein
C:Species: Mus musculus (house mouse)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30188
R:loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellison, Science 277, 232-235, 1997
A:Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis .
A:Reference number: Z20765, MUID:97362324; PMID:9211850
A:Accession: T30188
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1278 <LOF>
A:Cross-References: UNIPROT:O35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:G2251241;
C:Genetics:
A:Gene: Npc1
A:Map position: 18

Query Match 34.5%; Score 2385; DB 2; Length 1278;
Best Local Similarity 40.2%; Pred. No. 1.2e-158;
Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY	14	LLRLAQSEPTTTHQPGVCAFYDECKNPGLSGSLMTLSNVCLSNTARKITGDHLL 73
DB	12	LLLLCPAQVFSQ-----SCWYGECH---IATGD---KRYNCKYSGPPKPLPKDGYDL 59
QY	74	LQKICPLRYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSNQ 133
DB	60	VQELCPGLFF-DNVSLLCDIQLOQLKSNLQLPLQFLSRCPSCFYNLMTLFCELTCSHQ 118
QY	134	SLFINVTVAQIAGAGQLPA----VVAYRAFYQHSFAPQSYDSCSRVRVPAATLAVGTC 189
DB	119	SQFLNVTATEDYFDPKTPENKTNVKELEYVVCQSPANAMYNACRDVEAPSSNEKALGLLC 178
QY	190	GYVGSALCNAQRLNPFQDGTNGLAPLDI-----TFHLLPQAVGSGIQTPLNEGVARN 244
DB	179	GRDARA-CNATNWIEMFNKNGQAPFTIIPVFDLSIL-----GMEPRNRATKGN 229
QY	245	ESQGDVATCSQDCNASC-----PATARQALDSTP-----YLGQMPGSLV 286
DB	230	ESVDVTGFCSCQDCSIVCGPKPQPPMPMPWRWGLDMYIMVMTYVAFLVFVFGALL 289
QY	287	LIILICSVFAVVTILLVGRVAPARDKSMVDPKKGTSLSDKLSFTHTLLQFPQGWGT 346
DB	290	AVWCHRRYFVSEYTPIDSNIAFSVNS-----DKGEASCCDPLGAADFDDLRMTFKGA 345
QY	347	WVASHPLTILVLSVTPVVALAAGLVPTLTDPVELWSAPNSQARSEKAFHDQHGPPFR 406
DB	346	FCVRNPTCIIFFSLAFITVCSGLVFVQVTTNPVELWSAPHSEARLEKEYPDKHFGPPFR 405

Genes Dev. 10, 301-312, 1996
A;Title: Conservation of the hedgehog/patched signaling pathway from flies to mice: indu
A;Reference number: Z20752; MUID:96176226; PMID:8595881
A;Accession: T30172
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1434 <GO>
A;Cross-references: UNIPROT:Q61115; UNIPARC:UPI0000021C02; EMBL:U46155; NID:gl181884; PI
A;Experimental source: clone M2, M9
C;Superfamily: Drosophila membrane protein patched

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 9e-37; Mismatches 420; Indels 363; Gaps 44;
Matches 283; Conservative 171;

QY 356 LVLSPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHGFPPFTN-QVILTA 414
DB 88 LVUGLLIFGAFAGVLAANLETNVELWVEVGRVSRNELNYTRQKIGEEAMFQMLMIGT 147
QY 415 PNRSSRYDLSLLGPNKFGSGLDLDLLELE--LQERLRHQVWSPEAQRNLSODICY 472
DB 148 PKEEG-----ANVLTTEALLQHLDSALQASRVHVMYN---RQWLEHLICY 190
QY 473 AP-----LNPNTSLYDCCINSLLQYFQNNRTLLLTANQTLMOG----- 512
DB 191 KSGELITETGYMDQIIIEYLYPCLIIITPLDCFWEGAKLSQGT--YLLGKPLRLWTNFDPL 248
QY 513 -----TSQVDWKDHELY-----CAN-----APL----- 530
DB 249 EFLBELKKNYQVDSWEEMLNKAEVGHGYMDRPLNADPCDPATPNKNSTKPLDVALV 308
QY 531 -----TFKQGT-----ALALSCMADYGAPVFPFLAIGGYKKG 562
DB 309 LNGGCGQLSKYMHMQEELIVGTVKNATGKLVSAHALQTMQLMTPKQMYEHRGY--- 365
QY 563 DYSEBALMTSLNYPAGDPLQAQKLEWEAFLEEMRAFORRMAGMFQVTFIARSLE 622
DB 366 DY-----VSHINWNE-----DRAAILEAQRTYVEVHQSVAPNSTQKVLPTT-TTLD 414
QY 623 DSNRTTADLPFATSYIVIFLYISLALGSSYSSRVNVDKATGLGCVAVLGA VNA 682
DB 415 DILKSPDSVSVIRVAGVLLMAYACLTM---LRWD---CSKQCGAVGLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPEPREVHIGRALGR 742
DB 470 GLGLCSLIGISFNAAATTVLPFLALGVGDVDFLLAHAFSETGQNKRIIPFEDRTGELAKR 529
QY 743 VAPSMLLCSLSAICFFLQALTPMPAIVTFALTSLGLAVLDLFLQMSAFVALLSDSKQ 802
DB 530 TGASVALTSISNVTAFPMALPIPALRAFSLQAQVAVVVFNFAMVLLIPFALLSMDLYRR 589
QY 803 EASRLDVCCC-----VKQBEL-----PPGQGEGL----- 827
DB 590 EDRRLDIFCCTSPCVSRVQIPEQAYTEPHSNTRYSPPPPYTSHSFAHETHITWQSTVQ 649
QY 828 ----- 827
DB 650 LREYDPHTHYVTTAEPRSELSVQPVTVTDNLSCQSPSTSTEDLLSQSDSLHCL 709
QY 828 -----LLGFQKAYAPFLHWTIRGVLLFLFALFGVLSYMSCHISVGLDQELALPK 879
DB 710 BPPCTKWTLSPEAKHYAPFLKPKAKVAVVILLFLGLLGVSLYGTTRVRDGLDLDIVR 769
QY 880 DSVLLDYFLPLNRYEFGAPVFTVTLGVNFSSEAGMNAICSSAGCNFSPFKQIYAT- 938
DB 770 ETRYDFIAAQKPYFSF-----YNN-----YIVTQKADYPI 801
QY 939 -----EFPEQSYLAIPASS-----WVDDFDWL-----TPSS--- 965
DB 802 QHLLYDLHKSFSNVKVMLEENKQLPQMLHLFPRDMLQGLQDAFSDWETGRIMPNYKN 861
QY 966 -----CCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRPSVEQFHYKL-PWFLN 1017

DB 862 GSDDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGGINPSA--FYIYLTAWSN 915
QY 1018 D-----RPNIKCPKGGIAAY--STSVNLTSQGVLASRFMAYHKLKNSQD 1061
DB 916 DPAVAAASQANRPHRPEWHDK---ADYMPETRLRIPAAERPIEYAQFPYLLNGLRDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPQGTDPAFEPVPTTIVNVEQYVITLIPBGLMFLSLCL 1120
DB 973 FVEAIEKRVVICNNYVTSGLSSYPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
QY 1121 VPTAVSCLLGLDLSRGLNLLSIVMILVDTVGMALWDISYNAVSLINLVSAGVMSVE 1180
DB 1023 ACTFLLVCAVFLNPNWTAGII-VNVLALMTVELFGMGLIGIKLSAVPVVILIASVGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKETISMGSAGVAGVAMTNLPGILVLGLAKAQIQT 1238
DB 1082 FTVHVALAFLTAIGDKN---HRAMLALHMFAPVLDG-AVSTLLGVMLAGSEPDFIVRY 1137
QY 1239 PPRMLLTLLGLLHGLVPLPVLSYVG--PDVNPALAEQ---KRAEBAVAAMVVASCP 1293
DB 1138 PPAVLAILTVLGLNGLVLLPVLSPFGPCPEVSPANGNLNRLPTSPSPPPSVVRPAVPP 1197
QY 1294 NHPSRVS-TADNIYVNHSEPGSI-----KQAG 1319
DB 1198 GHTNNGSDSDSEYSSQTTVSGISELRQYEAQOQAG 1234

RESULT 6
T18538
patched protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18538
R;Marigo, V.; Scott, M.P.; Johnson, R.L.; Goodrich, L.V.; Tabin, C.
Development 122, 1225-1233, 1996
A;Title: Conservation in hedgehog signaling: induction of a chicken patched homolog by
A;Reference number: Z18958; MUID:96205046; PMID:8620849
A;Accession: T18538
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1442 <MAR>
A;Cross-references: UNIPROT:Q90693; UNIPARC:UPI00001328B4; EMBL:U40074; NID:gl335850; P
C;Genetics:
C;Superfamily: Drosophila membrane protein patched
C;Keywords: transmembrane protein

Query Match 9.1%; Score 628.5; DB 2; Length 1442;
Best Local Similarity 23.6%; Pred. No. 1.6e-35;
Matches 284; Conservative 165; Mismatches 421; Indels 335; Gaps 47;

QY 365 ALAAGLVFTELTPVELWSAPNSQARSEKAFHQHGFPPFTN-QVILTANRSTRYD 423
DB 111 APAVGLRAANLETNVELWVEVGRVSRNELNTRQKIGEEAMFQMLITQTPQEDG---- 166
QY 424 SLLGLPKNFGSLDLDLLELE--LQERLRHQVWSPEAQRNLSQDICYAP----- 474
DB 167 -----TNVLTTEALRQHLDSALQASRVHVMYN---RQWLEHLCYKSGELITEA 213
QY 475 --LNPNTSLYDCCINSLLQYFQNNRTLLLTANQTLMOG----- 512
DB 214 GYMDQIIIEYLYPCLIIITPLDCFWEGAKLSQGT--YLLGKPLQWINPDLPLBELKKI 271
QY 513 TSQVDWKDHELY-----CAN-----APLTF-KDGT-----ALALS--CMADY 546
DB 272 NYQVSEWEEMLNKAEVGHGYMDRPLNADPCDPATPNKNSTKPLDVALVSGG---Y 328
QY 547 GAPVFPF-----LAIGG-----YKGDYSEALIMTFSL-----NNYP----- 580
DB 329 GLSRKVMWQBELIIGTVKNSGKLVS-AQALQTMFQMLTPKQMYEHRGYEYVSHINW 387
QY 581 AGDPLQAQKLEWEAFLEEMRAFORRMAGMFQVTFIARSLEINRTTAEDLPFATSY 640

388	Db	NEDKAAALIEAWQRMVVEVVHQSVQNSQKVLSFTT--TTLDDILKSFSDSVIRVASGY 446
641	Qy	IVIFLYISIALGSSYSSWSRVMWDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILIQ 700
447	Db	LIJMLAVACLTM---LRWD--CAKSOQAVGLAGVLLVALSVAAAGLGLCSLIGISFNAATTQ 501
701	Qy	VVPFLVLSGADNIFIVLEBYQBLPRRPGPREVHIGRALGRVAPSMKLCISLSEAICPFL 760
502	Db	VLPFLALGVGDVDFLLAHAFSETGQWKRIFFEDRTGCEIKRTKGASVALTSISNVTAFFM 561
761	Qy	GALTPMPAVETPALTSGLAVIDLFLLOMSAFALLSIDSKRQBSASRLDVCCC-----812
562	Db	AALIPALRAFSLOAAVVVVFVAVVLLIPAILSMDLVRRDRDLFCFCTSPCVTR 621
813	Qy	---VKPO-----ELPPPGQGEGL-----827
622	Db	VIQEPQAYAENDNICYSSPPPYSSSHFAHETQITMQSTVOLRTEYDHPHTQAYTTAEPR 681
828	Qy	-----LLGFFQKAYAP 838
682	Db	SEISVQPVVTQDSLSCQSPESASSTRDLLSQFSDSSVHCLBPPCKWTLLSTPAEKHYAP 741
839	Qy	FLHLWITRGVWLLFLALFGVLSYSMCHISVGLDQELALPKDSVLDYFLFLNRYFEVGA 898
742	Db	FLKPKAKVWVIFLGLGLLSLYGTRVRDGLDLDIVPRDTRFYDFIAAQKYSF--799
899	Qy	PVYFVTTLGYNFSSEAGMNAICSSAGCNNSFTQIOYAT-----EPPEQSYLAI 948
800	Db	-----YNM-----YIVTQADYPNVQVHLLYELHRSFSNVTVILL 833
949	Qy	PA-----SSWYDDFIDWL-----TPSS-----CCRLYISGNKD 977
834	Db	EGDRQLPKWMLHYFRDMLQGLQADPDSDWETGKITYSNYKNGSDDAVLAYKLLVQTGNRA 893
978	Qy	KFCSTVNSLNCNCKMSITGWSVRPSVEQHKYL--PWFLND-----RPNI 1022
894	Db	K--PIDISQLT--KORLVADAGIINPNA--FYIYLTAWVNDPVAASQANTRPHRPEW 947
1023	Qy	KCPKGGGLAAY--STSVMLTSDGOVLASGRFMAYHKPLKNOSQDYETALRAABELAANITA-D 1079
948	Db	VHDK---ADYMPETRLRIPAAEPTEYQAQFPYINGLRETSDFVEAIEKVRACNNYTSLG 1004
1080	Qy	LKRVPGTDPAPEVPYITINVFYEQYITILPEGLFMLSLCLVPTFAVSCLLLLGLDLRSGL 1139
1005	Db	IASYPNG-----Ypf-----LFWEOYIGLRHWLLSISVVLACTFLVCALFLANPWTAGI 1054
1140	Qy	LNLSIVMILVDTVGFPMALNDISYNAVSLINLVSAGMSVEFVSHITSEF--AISTKPTW 1197
1055	Db	I--VVVALMTVELFGMMGLIGIKLSAVPEWVILIASVGIGVEFVHIALAFITAIGDKN--1111
1198	Qy	LERAKEATISGSAVFAGVAMTNLPGLIIVLGLAKAQITQIFFPRLNLIITLLGLHLGLVF 1257
1112	Db	--RRAVLALEHMFAPVLDG--AVSTLLGVNLAGSDFIVRVFFAVLAILIILGVNLGVL 1169
1258	Qy	LPVILSYVG--PDVNPALA---LEQKRAEEAAVAVMVVASCNPHSPRSTADNIYVNHSE 1312
1170	Db	LPVLLSPFGPYEVSFACGRNRLPTPSPPEPPIVRPALPGHTNNGS--DSSDSYSSQ 1227
1313	Qy	GSIGK 1317
1228	Db	TTVSG 1232

RESULT 7
S44797
F09G8.4 protein - *Caenorhabditis elegans*
C.Species: *Caenorhabditis elegans*
C.Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C.Accession: S44797
R.Anderson, K.
submitted to the EMBL Data Library, February 1993
A.Description: Sequence of the C. elegans cosmid F09G8.
A.Reference number: S44776

A:Accession: S44797
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-915 <AND>
A:Cross-references: UNIPROT:P34389; UNIPARC:UPI000013B859; EMBL:L11247; MID:g156280; P1L
C:Genetics:
A:Introns: 51/1; 214/3; 255/3; 326/3; 382/2; 409/3; 455/3; 505/3; 569/2; 702/3; 836/3

Query Match 8.7%; Score 600; DB 2; Length 915;
Best Local Similarity 22.7%; Pred. No. 8.9e-34;
Matches 236; Conservative 158; Mismatches 422; Indels 224; Gaps 37;

QY 10 LLWALLRLAQCSPYTTIHQPGYCAFYDCGKNPELSGLMSTLSNVSCISLNTPKITGD 69
DB 14 VLFLLLHLALCO-----AKVM-TECDGEDSNHPPCKTKNSTYLPITVTRSLNPT 64

QY 70 HLILLOKICPLRYTGN-TOACCSAQVLSLEASLITKALLTRCPACSDNFVNLCHNT 128
DB 65 YMARFEKYSYLVQEBEDKAQVCCTELQKGMTDRISNAATILGSCSPCFDNFACLWCQFT 124

QY 129 CSPNQSFLINVRVAQLGAQLPAVVAYAFYQHSFAEQSYDSCSRVRVPAATLAVGTM 188
DB 125 CSPDQSKFMKMET-----TQPKNVVVMKEKVNDRDFVEGLYESCRHTWPFANGALALMSL 180

QY 189 CGVYGSALCNAORWLNFGDTG-NGLA-----PLDITTHLEPQOAVGSGTQPLNEGVARCN 244
DB 181 GGRKVS-----FENFYGMGTKNLAQSIPINTEFQFSRMKNAMNIPITP-----CH 225

QY 245 ESQDDVATCSQDC---AASCPAIAARPOALDSTFVLGQMPGSLVLIIILCSVPVVVTL 301
DB 226 KSAQKVPACGAIDCPNHAQLVDISKVEHLGKTVFHPHP-DFEWLLKICGCLA-LTVL 283

QY 302 LVGF-----RVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFPQG-----WGTWVAS 350
DB 284 LVFTLKYSCHRRSAPNGEDCYVDLKGK-----LEVOPEGLCARYANAVIK 330

QY 351 WPLTILVLSVIPVVALAAG-LVFTELTTPDVELWAPNSQARSEKAF-HDQHGPPPTRN 408
DB 331 HPLIFVSLGLIVAAACCSGNFKSHLSHSDVQSAADGETTRNEKPIHS--FGPNHRIE 388

QY 409 QVILITAPNRSYRYSLLGPKNFGSLDLDLLELELOERLRLHQLVQVSPQAQRNLSLQ 468
DB 389 QIFINLPPT-----KSMFNPLPEEMFQLVGNIGNLT--ACYGNSSVKLD 432

QY 469 DICVAPLNPDNTSLYDCINSLLOYFQNNRTLLLTANTQITMGQTSQVDWKDHFLYCANA 528
DB 433 DICVKPIGKN-----HGCAINSPTNYFQ----- 455

QY 529 PLTFKQGTALALSCWADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQ 588
DB 456 -----YYTARTIMITVLIRG-----PE-DQ 474

QY 589 AKLWEEAFLEEMRAFORRMAGMFOVTTAERSLEDEINRTAED--LPIFATSVIVFLY 646
DB 475 AIAMETAFLNMWSRYEMKHAHF--TFWTETSVAAEIIHTAVETDKIVSVIACAAVLWVI 531

QY 647 ISLALGSY--SSWRGVVDKATLGLGGAIVLGVMAAMGFVSGLGRSSLVLIVQVVP 704
DB 532 TMLGNHWPESIIISALVHHKLLISISAVMISVISVCMISGIMFSLFGVHATDIAVLVLF 591

QY 705 LVLSVGADNIFIVLEYQRLPRRGPPE-----REVH--IGRALGRVAPSMLLCSISEATCF 758
DB 592 VITCLGINRIFVIRTFQANGCHYGLPNI SYRENMHRISNVMMRRSIPVLTNSLICSTCL 651

QY 759 FL-GALTP-----MPAVRTFALTSGLAVIDFLQMSAFVALLSDSKGQ----- 802
DB 652 FLAGGVLPVSVSPAVFVARHAGLAITMDTA FYLLVMLPLFQVDAREMSGKCEIWPW 711

QY 803 ----EASRLDCCVCKPQELPPQGGELLLGFQKAYAPFLHWHITRGVLLLFALFG 858
DB 712 YELSNESKINLCMEAVDGNLRSP-----VDWFKLAIPALLKKICRIWIATPFFVSLI 764

QY 859 VSLYSNMCHISVGLDQELALP-----KDSYLLDYFLPLNRYFEVGAPVYFT 904

```
Db 765 IACYCTLCLEFNGQMAESEVCIFFPSHLNGSFQSYLTUKFQNNENLNGPPLWV 824
Qy 905 TLGYNFSSBAGNAICSSAGCNFFSTQKIQ---YATEPPEQSILAIPASSWVDDFIDWL 961
Db 825 EGDVWKHDPKMQKFCCTLAGCDNSMGNKIRSLAYAEVY-KGNYLHGDVNIWLDVLOPM 883
Qy 962 TP-SGCCRLYISGPNKDKFC 980
Db 884 HPRGSCCKM-----DGKQFC 898

RESULT 8
T25600
hypothetical protein C32E8.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25600
R:Gattung, S.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C32E8.
A:Reference number: Z20056
A:Accession: T25600
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-933 <GAT>
A:Cross-references: UNIPROT:P91129; UNIPARC:UPI00000611CF; EMBL:U88308; PIDN:AA842325.1;
A:Experimental source: strain Bristol N2; clone C32E8
C:Genetics:
A:Gene: CESP:C32E8.8
A:Map position: 1
A:Introns: 145/1; 177/2; 385/1; 418/3; 833/3; 878/3

Query Match 8.7%; Score 598.5; DB 2; Length 933;
Best Local Similarity 22.9%; Pred. No. 1.2e-33;
Matches 251; Conservative 181; Mismatches 362; Indels 303; Gaps 37;

Qy 341 PGGWGTWASWPLTILVLSVIVPVALAAGLVFTELTTDPVELMSAPNSQARSEK-AFH-- 397
Db 14 FROLGLICDHPFPFVFPFLPFTAAMGVGLHNLPLSDAVLFTPLGAQSKMERMSIHEK 73

Qy 398 ----DQFGP---PFRNQVILTAPNRSYRVDLSLLGKPNFSGILDLDLLELELQER 450
Db 74 WPLTDNNYIPGRAVTSQSRIEQTALARN----DSNILDPKFANAVYQLDKY-----IQTR 124

Qy 451 LRHLQWSPQAQRNLSLODICVAPLNPDNTSLDYDCINSLLQYFQNNRTLLLTANQITLM 510
Db 125 VAVLH----- 129

Qy 511 GQTSQVDMKDHFLYCANAPLTPKDGTALE-----LSCMADYG-----APVFPFLAIGGY- 559
Db 130 -----NGHYYSYKNLCLQYKNGCPSNKHVHLSLDLHNGFNITYPYPRFSGEGVI 181

Qy 560 -----KGDYSE-----REALIMFSLNYP-----AGDPLAQAKLWEAFLEE 599
Db 182 GSSLGVTVMKGENETDILASAKAFMYIHLKHPHEMSYISGE-----WE---LEL 230

Qy 600 MRAFORMAGMF-QVTETAERSLEDEINRTTAEDLPFATSIVIVFLYISLALGSV--SS 656
Db 231 GMLTQYEDPDVISITYFHSQTLABELKKNADTLIPRFIISITLIVFSLCSLSDIGS 290

Qy 657 WSRVNVDSKATLGLGAVVAVLGAVMAMGFFSYLGIRSSLVITQVVPFLVSLVSGADNIFI 716
Db 291 PSIDVWLSKPIILSVISAGTALTGVGFLSLMGMPYN-DIVGVPFLVLAAGVDNMF 349

Qy 717 FVLEYQRLPRRGEPREVH--IGRALGRVAPSMILCSISEAICPFLGALTTPMVAVRTAL 774
Db 350 MVAAV-----RRTSRTHVHERMGECLADAASVILITSTSDVLSFGVGALTTPAVQIFCV 405

Qy 775 TSLGAVILDPLQMSAFVALLSLDSKQEAER-----LDVCCVK----- 814
Db 406 YTGVAIFAFYTIQITFFAACLALAMKHEASGRNSLFLTEAVSAEKKTSLSSTPQLFNLGS 465
```

```
Qy 815 -PQELPPPGQGGGLLPGFPQKAYAPFLHLHWITRGVVLLLP-LALFQVSLYSMCHISVGLD 872
Db 466 VPDHSHSDVKQPLTSRPFGEWYAPVLMHPVVRGIAWVFWIYLLGAS-YGCSRIKEGLE 524

Qy 873 QELALPKDSYLLDYFLFLNRYE-EVGAPVVFV-----TTLGYNPSSSE 913
Db 525 PWNLLVBDSYALPHYRLLEKIFWKYGOQVQIYINNAPDLRNHTSRDRVRVAMVLDPATSKH 584

Qy 914 A-GMNAICSSAGCNFFSTQKIQYATEPPEQ-----SYLAIPASSWVDDFIDWL 961
Db 585 AIGMESV-----QFWLPEMERYQKELEVOIIDSSFFYGLLHHFLASKTNPLAEDIYW- 637

Qy 962 TPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQFHKYLPWFLNDRN 1021
Db 638 -----GMPED----- 642

Qy 1022 IKCPKGLAAVSTSVNLSDGOVLAS-RFMAVHKPKNSQDYTEALRAARELANITADL 1080
Db 643 -----DDNGTWVKSFRPLGKMDLVTTMOQTDATMSFREVAARW----- 681

Qy 1081 RKPCTDPAFEVPPYITINVFYEQYLTLPEGLFMLSCLVPTPAVSCILLGLDLRSGLL 1140
Db 682 -----PEFNVTTPMPIWMTDQYIIIPNTVQNIILVMIIVIAVLFIQPMCS-LW 733

Qy 1141 NLLSYIMILVDVTFPMALWDISYNAVSLINLSVAGMSVEFVSHITRSPTKPTWLER 1200
Db 734 VALACASDFGVIQYITLWGLVNDIAISMTIITISGFSVDYSAHIAYGTVVSRDTPAQR 793

Qy 1201 AKEATISMGSAVFAGVAMTNLPGIILVGLAKAQLQIPEFRNLNLTLLGLAHGLVFLPV 1260
Db 794 VKEALSALGWPLSQG-AMSTIAVSVLADIPAYMI-VTFPKTVVLISLGLLHGLVFLPV 851

Qy 1261 ILSYV-----GPDVNPALALQKRAEEA-----VAAMVVASCPNHPSPSVS 1300
Db 852 LLSIFVPGCCIISSPHGHPHSAQKIEQIRINAISSSLDLRTVAPLRASSPISPHRLE 911

Qy 1301 -TADNIYNHSPFEGSIK 1316
Db 912 YDESPTVHNRSKNSIK 928

RESULT 9
T13952
membrane protein ptch2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T13952
R:Motoyama, J.; Takabatake, T.; Takeshima, K.; Hui, C.
Nature Genet. 18, 104-106, 1998
A:Title: Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog.
A:Reference number: Z17830; MUID:98122566; PMID:9462734
A:Accession: T13952
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1182 <MOT>
A:Cross-references: UNIPROT:O35595; UNIPARC:UPI0000001587; EMBL:AB010833; PIDN:BAA24691
A:Experimental source: strain BALB/cCrSlc
C:Genetics:
A:Gene: ptch2
C:Superfamily: Drosophila membrane protein patched
C:Keywords: transmembrane protein

Query Match 8.4%; Score 583.5; DB 2; Length 1182;
Best Local Similarity 24.2%; Pred. No. 1.8e-32;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

Qy 330 SFSTHTLLGQ-----PFQG-----WGTWVASWPLTILVLSVIVPVALAAGLVFTEL 376
Db 20 SSAPHILAGSLQAPLWLRAYFQGLLFLSGCRIQKHGKVLFLGLVAFGLALGLRVAVIE 79

Qy 377 TDPVELMSAPNSQARSEKAFHDQHPG-PFRFTNQVILTAPNRSYRVDLSLLGPKNFSGI 435
Db 80 TDLEQLWVEVGSVQELHYTKELGEEAAVTSQMLI-----QTAHQEGGNVLTPE----A 131
```



```
QY 996 ITWGSVRPSVEQPHKYLPHFLNDRPNKCPKGLAAYSTSVNLTSDQVL-ASRFMAYHK 1054
Db 630 -----VFLSQPOF-----AKYNRDVVLVTEDGEHLEASRI----- 658
QY 1055 PLK-----NSODVTEALRAARELAANITADLRKVPVGTDPAPFPPYITITNVFVEOYLTLIP 1110
Db 659 PVLHRHVGANGSRANRFLRLAE--TSELQ-----TGVYADPFO-----FAEQYNVAVLP 706
QY 1111 EGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIYMLVDTVGFPMALWDISYNAVSLIN 1170
Db 707 GTLSSIAVAGVAVVSLILPEPVAS-LWVSFISVINIGILGFMTFWSVRLDRFISMTV 765
QY 1171 LYSVAGMSVEFVSHITRFAISTKPTWLRKAETISMGSAPVAGVAMTNLPGLVILGLA 1230
Db 766 IVMSIGFCVDFAHLAYNFAKQNMDSRMRNALYAVGAPILMS-ATSTIIIGVSPMASA 824
QY 1231 KAQLIQIPFRLNLLITLGLHGLVFLPVILS--YVGPDPNPALALEOKRAEEAAV- 1287
Db 825 ESYVFR-SFKTIMVLVILGALHGLVFLVLSMFCGG-----SSKAKEHIDAVD 875
QY 1288 -MVASCPNHPSPRVST 1301
Db 876 QKLQAYNNPARTAS 890

RESULT 11
Ti8291
patched protein - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18291
R:Concordet, J.P.; Lewis, K.E.; Moore, J.W.; Goodrich, L.V.; Johnson, R.L.; Scott, M.P.;
Development 122, 2835-2846, 1996
A:Title: Spatial regulation of a zebrafish patched homologue reflects the roles of sonic
A:Reference number: Z18860; MUID:96379744; PMID:8787757
A:Accession: T18291
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1220 <CON>
C:Cross-references: UNIPROT:Q98864; UNIPARC:UPI00001328B3; EMBL:X98863; PIDN:CAA67386.1
C:Genetics:
A:Gene: Ptcl
C:Superfamily: Drosophila membrane protein patched

Query Match 8.3%; Score 572; DB 2; Length 1220;
Best Local Similarity 22.4%; Pred. No. 1-2e-31;
Matches 270; Conservative 174; Mismatches 431; Indels 328; Gaps 43;

QY 321 KGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSVIPVVALAAGLVFTELTTDPV 380
Db 51 KGVAVGQKAPLWIRARFQAFPLSLGCHIQRHCGKVLFIGLLVFGALSVGLRVAATETDIE 110
QY 381 ELWAPNSQARSEKAFHDHFG-PFRTHQVILTAPNRSSRYSDLSLLGPKNPSGILLDUD 439
Db 111 KLWVAGSRVSKELRYTKEQGEESVFTSQMLIQTPKQEG-----TNILTOE 157
QY 440 LLELLELERLRLHQLQVSPRAONISLQDICVAPLNP--DNT-----SLYDCINSL 490
Db 158 AL--LLHLEAALSASKVQSLVYKSWDLNKKICFKSGVPPIENVMTERMDKLPFCMVTP 215
QY 491 LQYFQNNRTL-----LLLTANO--TIMGQTSQVDWKDHFLYCA 526
Db 216 LDFWEGSKLOGSAVLPQMPDIPQWNLDPKLMEELSQTSLSEGFREMLDKAQGVHAYM 275
QY 527 NAPLTFKDTALALSCMADYDAP-----VFPFLAI---GGYKG----- 561
Db 276 NRPCLDPSDT-----DC--PHSAPNKPQWQVPNIAELQGGCGFSGKFMHQEELIGER 329
QY 562 -KD-----YSEBALINTFSL-----NNYPAGDPRLAQAKL-----WEBAFLRE 599
Db 330 VKDSQNALQSALQTMFLMSPKQLYEHFKDDYELHINWNEDKATALLESWQRKPFV 389
QY 600 MRAFA--QRRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISIALAGSYSSW 657
```

```
Db 390 VHGISIPONSSSNVYAFSTT---TLNDIMKSFSDSVIRVAGGYLLMLAYACVTM---LRW 443
QY 658 SRMVVDSKATLGLGAVVVLGAVMAAMGPFYSILRSLSILQVVPFLVLSGADNIFIF 717
Db 444 D--CAKSGQAVGLAGVLLVALSVAAGLGLCSLLGSFNAATTVQLVPSLAGIGVDDMFL- 500
QY 718 VLEYQRLPRRPGEPREVIHGRALGRVAPSMMLCSISEAICPFLGALTMPPAVRTFALTSG 777
Db 501 -LGHSTFETRSNIPKERTGDCRLRTGTSTVALTSVNMIAFFMAALVPIPALRAPSLQAA 559
QY 778 LAVIILDFLOMSAPFALLSILDSKROBASRLDVCCC-----VKQOE----- 817
Db 560 VVVVFNFAMALIFPAILSLDLRREDKELDLCCFPSPCSSRVTIQIQQLSANDNDHQ 619
QY 818 -----LPPPGQ----- 823
Db 620 RAPATPTTGTSTITTTTHTTTVQAPTQCDAAAGQHIVTLPPTTSQISLTPPSPMVLSTPTP 679
QY 824 -----GEGE-----LLGFFQKAYAPFLLHWITR 846
Db 680 TTDPYGSQVFTTSSSTRDLLAQVBERPKGREGCVLPFPFRNLLSSPAREKYAPLLKPEYK 739
QY 847 GVLLLLFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAPVYFVTTL 906
Db 740 TVVVVVFALLSLSLYGTTTWHVDGLYLTDIVPRDTQVEYEFITAQFKYFSF-YNMVLVTMD 798
QY 907 GYNFSSEAGMNAICSSAGCNCNPFQTKQYATFEPQSVLA-----IPASSWDDPDIM 960
Db 799 GDYAR-----SQRLQLQHNAFNSVKYVVDGNHKLGP-RMWLHYFQDM 841
QY 961 LTPSSCC-----RLYISGPNKDKFCPSTVNSLNCNKC 993
Db 842 LKGLQATPDADWEAGKITVDSYRNGTGDGALAYKPLIQTSKKEPNYSQLSRRLVDG- 900
QY 994 MSITWGSVRPSVEQPHKYL-PWFLNDRPNKCPKGLAA-----YST----- 1034
Db 901 -----DGLIPPEV--FYIYLTVMVNSD-----PLGYAASQANFYPHPREWIHDKYDTTGE 948
QY 1035 SVNLTSDGQVLASRRPMAYHKPLKNSQDYTEALRAARELAAN-ITADLRKVPGTDPAPFVE 1093
Db 949 NLRIAPAEPLPEPAQPFYVINGLURQASDFIEATESVRTICEEFMRQIKYKNG-----Y 1002
QY 1094 PYTITNVFVEOYLTLIPBGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIYMLVDTV 1153
Db 1003 PP-----LFWEQYIGLRHWFLLSISVVLACTFLVCAILLNLPWTAGVI-VFILPMVTVELP 1057
QY 1154 GPMALWDISYNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLRKAETISMGS 1211
Db 1058 GIMGLIGIKLSAIPVIVILLIASVGIVFTVHIALGFLTAIGDRNT---RSAVAMEHMFAP 1114
QY 1212 VPAGVAMTNLPGLVILGLAKAQLIQIPFRLNLLITLGLHGLVFLPVILSVGP--DV 1269
Db 1115 VTDG-AISTLLGLVLMAGSEDFIMRYFFAVAILTLTGILNGLVLLPVLMSLGMPPAEV 1173
QY 1270 NPA 1272
Db 1174 VPA 1176

RESULT 12
T29590
hypothetical protein F55P8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29590
R:Gattung, S.; Scheet, P.; Kemp, K.
A:Description: The sequence of C. elegans cosmid F55P8.
A:Reference number: Z20647
A:Accession: T29590
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
```

A;Residues: 1-889 <GAT>
A;Cross-references: UNIPROT:P91346; UNIPARC:UPI000017BA20; EMBL:U80447; PIDN:AAB37812.1;
A;Experimental source: strain Bristol N2; clone F55F8
C;Genetics:
A;Gene: CESP:F55F8.1
A;Map position: 1
A;Introns: 36/3; 66/3; 98/3; 149/2; 200/3; 240/2; 279/3; 358/1; 481/2; 569/3; 606/3; 697/3

Query Match 7.8%; Score 539.5; DB 2; Length 889;
Best Local Similarity 23.0%; Pred. No. 1.5e-29;
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;

QY 401 FGPFRFTNQVIL-----TAPNRSSVRYDLSLLGPKNFSGILDLILLLELLE 446
DB 31 FGPYSYTERRIIHDAPLVDGTFVAGRAVTS--REVQVAVVARSGGNLDVRFNSNELK 88
QY 447 LOERLRHLQVNSPEAQRNLSLODICVAPLNPNTSLYDCCINSLLOVYFQNNRTLLLTAN 506
DB 89 MESFIRN-NITVQFSNRWTFADLCIA--GPDGR-----CANNDHIQ----- 127
QY 507 QTLMGQTSQVDWKDHFYLCANAPLTFKDGTTALALSCMADYGAPVPPFLAIGYK----- 560
DB 128 --LASRLHQ-----HGINTITYFTVRLSDKSAYIAS-----ALGGVKLAKGDN 167
QY 561 GKD-YSEAEALIMTSLNNY-----AGDPRLAQAKLWEAPLEEMRAFORMAGMFQV 613
DB 168 GENIIVEATAMLLIYQLKFPNEISYVSG-----LWEREFKNKMDY-KQAKYIS 218
QY 614 TTAERSLEDEINRTAEDLPATSYIVILY-----ISLALGYSWSRSMVMDSKAT 667
DB 219 TYFHSQTLSDLRNARNELAPKIFAGFVILCVFSLCSIVTIKSGYIDW---VVTXPI 274
QY 668 LGLGGVAVVLGAVMAAGFFSYLGIRSSILVILQVVPFLVSLVGADNIFIVLEYQRLPRR 727
DB 275 LSVLGVSNAGMGASANGMLTYLEIQYN-DIIAVMPFLVAVGTDMFMELVASLKRTR- 332
QY 728 PQEPREHVHTRALGRVAPNMLCSLSEALCFPLGALTMPAVRTALTSLGLAVILDFLQ 787
DB 333 -NLKYDQRIAEACMADAASILITALTALDTSFGVTGTTTTPAVQIFCIYTMCAILLTFAVQ 391
QY 788 MSFAVALLSDSKQE-----ASRLDV---CCVKPQBLPPPGOG--- 824
DB 392 LTFECAILVYTRIEREQGLHSWLBRPAVYTSSTPLNVLFWLGSQPK-PLPSCGTVSS 450
QY 825 -----EGLLGFQKAVAPFLH-WITRGVLLFLALFGVSLYS 863
DB 451 TSSVSTMTSQATSPASKHLHCAATSFFRNWYAPVLMQPMI-RAIAGLWLYLIGISYIG 509
QY 864 MCHISVGLDQELALPKDSYLLDYFLNLR-YFEVGAPVYFVTI-----LGYNFSSE 913
DB 510 CTHLKEGLEPANILLVDSYATPHYRLEKHYWHYGASLQIVVSNPPDLRDPVERINMDM 569
QY 914 AGMAICSSA-GCNNFSF-----TOKIQYATE---FPEQSYLAIPASSWYDDFI 958
DB 570 ASTFANCKVAIGDSQVFWLREMQVSEIHKIQYDNEKFYDHAAQYIYSDMSQPWVVDV 629
QY 959 DWLTPSSCCRLYISGPNKDKFCPSTVNSLNCNMSITMGSVRPSVEQPHKYLPLFLND 1018
DB 630 -----WGRNN 634
QY 1019 RPNIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTRALRAARELANITA 1078
DB 635 -----NSERIKTFRPMIGMRDISTTTQTEATNTFRFIAISRF-- 672
QY 1079 DLKRVPGTDPAFEVFPYITNVFEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
DB 673 -----BOYNVTTTWPLWFLFDQVALVPNTMQDIIIVAVCMVLISALLIPQPCVSF 723
QY 1139 LLNLLSIVMLVDVTFGMALWDISYNAVSLINLVSAGMVSVEFVSHITRSFALSTPTWL 1198
DB 724 WV-AVTIGSIDLGVLFMTLNNVNLDAISMITIIMSFGVSVDYSAHITYAVYSKESST 782
QY 1199 ERAKEATISGSAVFAGVAMTNLPGILVLGLAKQLIQIFFRNLNLLITLLGLHLGLVFL 1258

DB 783 ARVCDALGDLGWPAQG-AMSTILAVSVLSVDVPAYMI-VTFEKTVELAISIGFLHGLVFL 840
QY 1259 PVILS-YVG 1266
DB 841 PLMLSVPVG 849

RESULT 13
T27969
hypothetical protein ZK675.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T27969
R;Sims, M.
submitted to the EMBL Data Library, November 1994
A;Reference number: Z20448
A;Accession: T27969
A;Status: preliminary; translated from GE/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1405 <WIL>
A;Cross-references: UNIPROT:Q09614; UNIPARC:UPI00000835A5; EMBL:Z46812; PIDN:CAA86843.1;
A;Experimental source: clone ZK675
C;Genetics:
A;Gene: CESP:ZK675.1
A;Map position: 2
A;Introns: 31/1; 59/3; 100/3; 244/2; 485/1; 950/2; 1298/2; 1335/3
C;Superfamily: Drosophila membrane protein patched

Query Match 7.4%; Score 514; DB 2; Length 1405;
Best Local Similarity 20.2%; Pred. No. 1.7e-27;
Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLSDKLSTHTLLGQFPQCGWTWY--ASWPLTILVLSVIPVWALAAGLVPTELTD 378
DB 100 RKGATGNRYALYSRLIQKLLFALGNTVHRNWSI-ILAVSNIFAVC-CYGLQYVHIETD 157
QY 379 PVELWSA-----PNSQARSEKAFHDQFGPPF-RTN-----QVILTAPNRSS 419
DB 158 IVKLWVAOGGRDLDELNPLFNKEAMRNVTCDS--GPELPRENGLGGGVQVLIQTPEYEG 215
QY 420 YRYDSLLGPKNFSGILDLDLLELLELLEQLRHLQVNSPEAQRNI-----SLQDICY 472
DB 216 --QDALAAGP-----LLKHVEIMKHIASEFNVSVHGVDMWSLDCIF 253
QY 473 APLNPD-----NTSLYD-----CC-INSILQYFQNNRTL-----LILL 503
DB 254 KPAPSPVAADSAASSLGVDIKIVPCIWITPDCFEWSKALGPHPSLPKSSLGLPLGMLL 313
QY 504 TANQTLMGQTSQVDWKD-----HFLYCANAPLTP----- 532
DB 314 SS-----LSDGDMIRWSDPDPIAVIDEIHRSFNLGSHYTPFERAGVSHGYMDRPFCDPLDP 369
QY 533 -----KGTALALSCHADYGAPVFPFLAIGYKGYKDYSEAE 568
DB 370 ECPFMAKNYFDVCPHIDRVREIAKYGTELEBEKKKSGYSFFDEL---GRKKREAGDQ 426
QY 569 ALI----- 571
DB 427 KMIHPAQPADSIPTTIEDVAPQVPYSTAPIPTTTTILSPEEARAAEKEKKQKARELKDYC 486
QY 572 -----MTFSLNNYP-----AGDP 584
DB 487 KSVKRSAPFWLKKKNKQKWPVEVNSNNYQNVDYAAEMTGCGSGFASNVNLNWPEDMTLGNP 546
QY 585 RLQA--AKL-----NEEAPLES-MRA 602
DB 547 RRKAKGKGLSGDALQSFLVASPADVFLRFKQKPGRNSNMTKGLDMDAWNETHAAEQVLQA 606
QY 603 FORRMAGMQVFTFAERSL-----EDEINRTTAEDLPIFATS-----YIV 642
DB 607 WQR-----NFT--KSLYNHKNVDEDEGNERETLH--PLASTSIADMLEBFCQFNYYI 654


```

Db      857 MVTIDIGVIGFLSLWSVKLDPISMITIIMSGSIFBSAHITHGFVNSDNLSPAFDRCDV 916
Qy      1204 ATISGSAVFAVAMTNLPGILVLGLAKAQIQIFFFFRLNLLITLLGLHGLVFLPVLS 1263
Db      917 AMEKLANPVVG-SLSTILGVFLAPIDSVMVLFKTSISLV-LIIGAWHALMLPILLS 974
Qy      1264 YVGP-----DVNPALALEQKRAEAAVAVMVASCP 1293
Db      975 MCIPIVIERLSASKASDRRRKLNKNSVAINLP 1010

RESULT 15
S06119
membrane protein patched - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: S06119; A33468
R:Nakano, Y.; Guerrero, I.; Hidalgo, A.; Taylor, A.; Whittle, J.R.S.; Ingham, P.W.
Nature 341, 508-513, 1989
A:Title: A protein with several possible membrane-spanning domains encoded by the Drosophila
A:Reference number: S06119; MUID:90015164; PMID:2797178
A:Accession: S06119
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1299 <NAK>
A:Cross-references: UNIPROT:P18502; UNIPARC:UPI000014EA18; GB:X17558; NID:98389; PIDN:CA
R:Hooper, J.E.; Scott, M.P.
Cell 59, 751-765, 1989
A:Title: The Drosophila patched gene encodes a putative membrane protein required for se
A:Reference number: A33468; MUID:90058658; PMID:2582494
A:Accession: A33468
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-110,'R',112-273,'G',275-331,'R',333-635,'P',637-861,'DVF',878,'Y',880-1299
A:Cross-references: UNIPARC:UPI000016BCA3; GB:M28418; GB:M28999; NID:g552097; PID:g55209
C:Genetics:
A:Gene: FlyBase:ptc
A:Cross-references: FlyBase:FBgn0003892
A:Map position: 2 44D3-D4
C:Superfamily: Drosophila membrane protein patched
C:Keywords: glycoprotein; transmembrane protein
F;74-92/Domain: transmembrane #status predicted <TM01>
F;427-448/Domain: transmembrane #status predicted <TM02>
F;456-503/Domain: transmembrane #status predicted <TM03>
F;529-555/Domain: transmembrane #status predicted <TM04>
F;557-585/Domain: transmembrane #status predicted <TM05>
F;677-699/Domain: transmembrane #status predicted <TM06>
F;967-1017/Domain: transmembrane #status predicted <TM07>
F;1019-1047/Domain: transmembrane #status predicted <TM08>
F;1061-1086/Domain: transmembrane #status predicted <TM09>
F;1093-1121/Domain: transmembrane #status predicted <TM10>
F;142,298,335,388,807,861,1194,1271/Binding site: carbohydrate (Asn) (covalent) #status

Query Match      6.78; Score 464.5; DB 1; Length 1299;
Best Local Similarity 20.48; Pred. No. 4.4e-24;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

Qy      321 KGTSLSKLSFSTHTLGLQFGQGTWVASWPLTILVLSVIPVVALAAGLVFTELTDPV 380
Db      43  KGKARGSRITAIYLSRFQSHLETGLTSSVQKHAGKLVFVAILVLSTFCVGLKSAQIHSKVH 102
Qy      381 ELWSAPNSQARSEKAFHQHFG-PFFETNQVILTA---PNRSSRYVDSLLGPKNFSGIL 436
Db      103 QLWITQEGGLEAEIAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
Qy      437 DLDLLLELLEQLERLRLQVMSPEAQRNISLQDICYAPLNPNTSLY-----DCC 486
Db      154 HLEVLVATAVKVLHYDTE-W-----GLRDCNMPSTPFGIYIQLRLHLPIS 204
Qy      487 INSLQVFN-----NRTLLLTAN-----QTLMGQTSQVDKDH 521
Db      205 IITPLDCFEWGSQLLGPESAVVIGLNLQRLWTLTINPASNQYMKQKSEKISFDFETV 264

```

```

Qy      522 FLYCANAPLTFKDGTTALALSCM-----ADYGAPVPFPFLAIG--GYK 560
Db      265 EQYWKRAATA---SGWKEPCLNPLNPNCPDTAPNKNSTOPPDVGA-----ILSGGCGYA 317
Qy      561 GKDYSEAEALIMTFSLNPNYPAGDPLAQA-----589
Db      318 AKHMWPEELIVGAKGN-RSGHLRKAQALQSVMQLMTEKEMYDQMDNYKVHHLGWTOE 376
Qy      590 -----KLWEEAF---LEENRAFQRWAGMAGFQVFTFAERSLEDEINRTTAEDLPPIPATSY 640
Db      377 KAAEVLNAWQRFNSREVEQLLRKQSRATNYDVIYVSSAALDDILAKFSPHSALSIVIGV 436
Qy      641 IVIFLYISLALGSYSSMSRVWDKATLGLGGVAVVLGAVMAAGFFSYLGISSVLVQ 700
Db      437 AVTVLY---AFCTLLRW-RDPVRGQSSGVAGVLLCMFCSTAAGLGLSALLGIVFNAASTQ 492
Qy      701 VVPLVLISVAGADNIFIVLEYQRLPRPGBPREVHIGRALGRVAPSMMLCSLSAICFFL 760
Db      493 VVPLALGLGVDFHIFMLTAAYASNNR--EQTKL-----ILKKVGPSILFSACSTAGSFFA 546
Qy      761 GALTMPAVRTFALTSLGLAVLLDQLQMSAFVALLSLDSKQEAASRLDV-CCCV-----K 814
Db      547 AAFIVPALVKVFCLOAAIVMCSNLAALLVPPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606
Qy      815 PQELPP-----PGQEGEGL-----LIGFF 832
Db      607 PKVAPPVPLNNGNGRGARHPKSCNNRVALPAQNPLEQRADIPGSSHSLSASFSLATFA 666
Qy      833 QKAYAPFLHWTITGVVLLLFALFGVLSYMSCHISVGLDQELALPKDSYLLDFELFLNR 892
Db      667 FQHYTFPLMSWVKFVTWGFALAALISLYASTRLQDGLDIIDLVPKDSNBHKLDAQTR 726
Qy      893 YEEVGAPVYVTTTLYGYNFSSEAGNAICSSAGCANNFSFTOKIQTATPEPQSYLAIP--- 949
Db      727 LF--GFTSYMYAVTQ-----NFEPYTOQLLRDY-HDSFVRVPHVI 764
Qy      950 -----ASSWVDDFDIWLTPSSCRLYISGPNKDKFCFSTVNSLNCUNK----- 993
Db      765 KNDNGGLPDFWLLLFSEWL-----GNLQKIPDEEYRDGRITKECFWFPNASSDA 812
Qy      994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db      813 ILAYKLIYQGTGHVDPNPVDKELVLTNRLVNSDGIINQRAFYNLSAW-----ATNASSPTL 868
Qy      1029 LAA--YSTSVNLTSDGOVLASRFMAYHKPLKNSQD-----YTEALRAAR 1070
Db      869 LRANCIRNRANGASQCKLYPEPRQYFHP--NEYDLKIPKSLPLVYAQMPPYLHGLTDT 926
Qy      1071 ELAANITADLRKVPCTDPAFEV--PPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSC 1128
Db      927 QI-KTLIGHIRDLSVKYEGFGLPNYPSPGIPFIFMEQYMT-LRSSLAMTILACVLLAALV 984
Qy      1129 LLLGLDLSGLNLLISVIMLVDTYGFALWDISYNAVSLINLVSAGVMSVFEVSHITS 1188
Db      985 SLLLSLWAAVILSVLASLAQIFGAMTLLGIKLSAIPAVILISVGMMLCF--NVLLS 1042
Qy      1189 FAISTKPTWLERAKATISMGSAVPAVAMTNLPGILVLGLAKAQIQLIFFRLNLLITL 1248
Db      1043 LGFMTSVGNRRQVQLSQMSLGPLVHGLMTSGVAVFMLSTSPFEFVIRHFCWLLILVLC 1102
Qy      1249 LGLHGLVFLPVLSYVGVDPVNPALALEQKRAEAAVAVMVASCPHPSRVSTADNIYN 1308
Db      1103 VGACNSLLVFPILLSVVGPEAE-LVPLE-----HPDRISTPSPPLPVR 1143
Qy      1309 HS 1310
Db      1144 SS 1145

```

Search completed: April 7, 2006, 19:22:06
Job time : 63 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: April 7, 2006, 19:13:14 ; Search time 265 Seconds
(without alignments)
3546.278 Million cell updates/sec
Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGATSNFLPNNRGQF 1332
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	6872.5	99.5	1359	1	NPCL1_HUMAN	Q9uhc9 homo sapien
2	5421.5	78.5	1331	1	NPCL1_RAT	Q6t3u3 rattus norv
3	5407	78.3	1333	1	NPCL1_MOUSE	Q6t3u4 mus musculus
4	3179	46.0	1132	2	Q4T749_TETNG	Q4t749 tetraodon n
5	2413	34.9	1277	2	Q9JUG3_CRIGR	Q9j1g3 cricetus
6	2412	34.9	1277	1	NPCL1_PIG	P56941 sus scrofa
7	2405	34.8	1276	2	Q8MKD8_FELCA	Q8mkd8 felis silve
8	2404.5	34.8	1276	2	Q9GK52_CANFA	Q9gk52 canis fami
9	2402.5	34.8	1278	1	NPCL1_HUMAN	O15118 homo sapien
10	2402	34.8	1289	2	Q59GR1_HUMAN	Q59gr1 homo sapien
11	2400	34.7	1277	2	Q7TMD4_MOUSE	Q7tmd4 mus musculus
12	2399	34.7	1276	2	Q9NQO0_FELCA	Q9nqo0 felis silve
13	2395	34.7	1276	2	Q8MI49_FELCA	Q8mi49 felis silve
14	2385	34.5	1278	1	NPCL1_MOUSE	O35604 mus musculus
15	2367.5	34.3	1277	2	Q9GJC9_BOVIN	Q9gjc9 bos taurus
16	2361	34.2	1286	2	Q9TT75_RABIT	Q9tt75 oryctolagus
17	2338	33.8	1209	2	Q4RWY5_TETNG	Q4rwy5 tetraodon n
18	2262	32.7	1287	2	Q9VL24_DROME	Q9vl24 drosophila
19	2256	32.7	1287	2	Q9U5W1_DROME	Q9u5w1 drosophila
20	2253	32.6	1287	2	Q7YU59_DROME	Q7yu59 drosophila
21	2205	31.9	1291	2	Q7Q409_ANOGA	Q7q409 anopheles g
22	1864.5	27.0	1223	2	Q9VRC9_DROME	Q9vrc9 drosophila
23	1760	25.5	1003	2	Q7PS03_ANOGA	Q7ps03 anopheles g
24	1725	25.0	1275	2	Q51NK7_MAGGR	Q51nk7 magnaporthe
25	1712.5	24.8	1361	2	Q7XUB7_ORYSA	Q7xub7 oryza sativ
26	1700.5	24.6	1271	2	Q5BBG1_EMENI	Q5bbg1 aspergillus
27	1659.5	24.0	1330	2	Q5KGS9_CRYNE	Q5kgs9 cryptococcu
28	1658.5	24.0	1330	2	Q5S8D4_CRYNE	Q5sd84 cryptococcu
29	1643	23.8	1273	2	Q4WNG5_ASEFU	Q4wng5 aspergillus
30	1614.5	23.4	1295	2	Q41OK4_GIBZE	Q41ok4 gibberella
31	1549	22.4	1264	2	Q6BT03_DEBHA	Q6bt03 debaryomyce

32 1527.5 22.1 1275 2 Q9SHN9_ARATH
33 1511.5 21.9 1342 2 Q9TVK6_DICDI
34 1508 21.8 1239 2 Q6CBA1_YARLI
35 1500 21.7 1162 2 Q7RWL9_NEUCR
36 1461.5 21.2 1256 2 Q59VW0_CANAL
37 1452.5 21.0 1397 2 Q5S1C5_DICDI
38 1437.5 20.8 1489 2 Q4PEB3_USTMA
39 1394 20.2 1055 2 Q9SVF0_ARATH
40 1386.5 20.1 620 2 Q4H344_CIOIN
41 1359 19.7 1382 2 Q618V4_CABBR
42 1341.5 19.4 1383 1 NPCL1_CHEEL
43 1329 19.2 1170 2 Q12200_YEAST
44 1323 19.1 1339 2 Q5ORB6_ENTHI
45 1312.5 19.0 1178 2 Q750G1_ASHGO

ALIGNMENTS

RESULT 1
NPCL1_HUMAN
ID NPCL1_HUMAN STANDARD; PRT; 1359 AA.
AC Q9UHC9; Q6R3Q4; Q9UHC8;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Niemann-Pick C1-like protein 1 precursor.
GN Name=NPCL1;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 2 AND 3), AND TISSUE SPECIFICITY.
RX MEDLINE=20247253; PubMed=10783261; DOI=10.1006/geno.2000.6151;
RA Davies J.P., Levy B., Ioannou Y.A.;
RT "Evidence for a Niemann-Pick C (NPC) gene family: identification and characterization of NPC1L1";
RL Genomics 65:137-145(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), AND TISSUE SPECIFICITY.
RX PubMed=14976318; DOI=10.1126/science.1093131;
RA Altmann S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
RA Tetzeloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
RA Murgolo N., Graziano M.P.;
RT "Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol absorption";
RL Science 303:1201-1204(2004).
RN [3]
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15671032; DOI=10.1074/jbc.M409110200;
RA Davies J.P., Scott C., Oishi K., Liapis A., Ioannou Y.A.;
RT "Inactivation of NPC1L1 causes multiple lipid transport defects and protects against diet-induced hypercholesterolemia";
RL J. Biol. Chem. 280:12710-12720(2005).
RN [4]
RP INDUCTION.
RX PubMed=15604518; DOI=10.1194/jlr.M400400-JLR200;
RA van der Veen J.N., Kruit J.K., Havinga R., Baller J.F.W., Chimini G.,
RA Lestavel S., Staels B., Groot P.H.E., Groen A.K., Kuipers F.;
RT "Reduced cholesterol absorption upon PPARdelta activation coincides with decreased intestinal expression of NPC1L1";
RL J. Lipid Res. 46:526-534(2005).
RN [5]
RP FUNCTION.
RX PubMed=15928087; DOI=10.1073/pnas.0500269102;
RA Garcia-Calvo M., Lissack J., Bull H.G., Hawes B.E., Burnett D.A.,
RA Braun M.P., Crona J.H., Davis H.R. Jr., Dean D.C., Detmers P.A.,
RA Graziano M.P., Hughes M., MacIntyre D.E., Ogawa A., O'Neill K.A.,
RA Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,
RA Ujjainwalla F., Altmann S.W., Chapman K.T., Thornberry N.A.;

RT "The target of ezetimibe is Niemann-Pick C1-Like 1 (NPC1L1).";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137 (2005).
 RN [6]
 RP VARIANTS LEU-55 AND ASN-1233.
 RX PubMed:15679830; DOI=10.1111/j.1399-0004.2004.00388.x;
 RA Wang J., Williams C.M., Hegele R.A.;
 RT "Compound heterozygosity for two non-synonymous polymorphisms in
 RL NPC1L1 in a non-responder to ezetimibe.";
 RL Clin. Genet. 67:175-177 (2005).
 CC -1- FUNCTION: Play a major role in cholesterol homeostasis. Is
 CC critical for the uptake of cholesterol across the plasma membrane
 CC of the intestinal enterocyte. Is the direct molecular target of
 CC ezetimibe, a drug that inhibits cholesterol absorption. Lack of
 CC activity leads to multiple lipid transport defects. The protein
 CC may have a function in the transport of multiple lipids and their
 CC homeostasis, and may play a critical role in regulating lipid
 CC metabolism.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation
 CC of brush border membranes from proximal enterocytes suggests
 CC considerable association with the apical membrane fraction. Exists
 CC as a predominantly cell surface membrane expressed protein (By
 CC similarity). According to Ref.3 localizes in a subcellular
 CC vesicular compartment rich in RAB5.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9UHC9-1; Sequence=Displayed;
 CC Name=2; Synonyms=NPC1L1DELTA15;
 CC IsoId=Q9UHC9-2; Sequence=VSP_015314;
 CC Name=3; Synonyms=NPC1L1T;
 CC IsoId=Q9UHC9-3; Sequence=VSP_015312, VSP_015313;
 CC TISSUE SPECIFICITY: Widely expressed. Expressed in liver. Also
 CC expressed in small intestine, pancreas, kidney, lung, pancreas,
 CC spleen, heart, gall bladder, brain, testis, stomach and muscle.
 CC -1- INDUCTION: Expression is decreased in Caco-2 cells upon PPAR
 CC activation.
 CC -1- PTM: Highly glycosylated (By similarity).
 CC -1- POLYMORPHISM: Variations in NPC1L1 gene could be associated with
 CC non-response to ezetimibe treatment.
 CC -1- SIMILARITY: Belongs to the patched family.
 CC -1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF192522; AAF20396.1; -; mRNA.
 CC EMBL; AF192523; AAF20397.1; -; mRNA.
 CC EMBL; AY515256; AAS56939.1; -; mRNA.
 CC EMBL; AY437865; AAR97886.1; -; mRNA.
 CC Ensembl; ENSG0000015520; Homo sapiens.
 CC HGNC; HGNC:7898; NPC1L1.
 CC MIM; 608010; -;
 CC GO; GO:0015018; P:sterol transport; TAS.
 CC InterPro; IPR004765; NP_C_type.
 CC InterPro; IPR003392; Patched.
 CC InterPro; IPR000731; SSD_5TM.
 CC Pfam; PF02460; Patched; 1.
 CC TIGRfams; TIGR00917; 2A060601; 1.
 CC PROSITE; PS50156; SSD; 1.
 CC Alternative splicing; Cholesterol metabolism; Glycoprotein;
 CC Lipid metabolism; Polymorphism; Signal; Steroid metabolism;
 CC Transmembrane.
 CC SIGNAL 1 21 Potential.
 CC CHAIN 22 1359 Niemann-Pick C1-like protein 1.
 CC TOPO_DOM 22 284 Extracellular (Potential).
 CC TRANSMEM 285 305 1 (Potential).
 CC TOPO_DOM 306 351 Cytoplasmic (Potential).
 CC TRANSMEM 352 372 2 (Potential).
 CC TOPO_DOM 373 632 Extracellular (Potential).
 CC TRANSMEM 633 653 3 (Potential).

FT	TOPO_DOM	654	666	Cytoplasmic (Potential).
FT	TRANSMEM	667	687	4 (Potential).
FT	TOPO_DOM	688	696	Extracellular (Potential).
FT	TRANSMEM	697	717	5 (Potential).
FT	TOPO_DOM	718	742	Cytoplasmic (Potential).
FT	TRANSMEM	743	763	6 (Potential).
FT	TOPO_DOM	764	776	Extracellular (Potential).
FT	TRANSMEM	777	797	7 (Potential).
FT	TOPO_DOM	798	846	Cytoplasmic (Potential).
FT	TRANSMEM	847	867	8 (Potential).
FT	TOPO_DOM	868	882	Extracellular (Potential).
FT	TRANSMEM	883	903	9 (Potential).
FT	TOPO_DOM	904	1139	Cytoplasmic (Potential).
FT	TRANSMEM	1140	1160	10 (Potential).
FT	TOPO_DOM	1161	1168	Extracellular (Potential).
FT	TRANSMEM	1169	1189	11 (Potential).
FT	TOPO_DOM	1190	1191	Cytoplasmic (Potential).
FT	TRANSMEM	1192	1212	12 (Potential).
FT	TOPO_DOM	1213	1236	Extracellular (Potential).
FT	TRANSMEM	1237	1257	13 (Potential).
FT	TOPO_DOM	1258	1359	Cytoplasmic (Potential).
FT	DOMAIN	632	797	SSD.
FT	COMPBIAS	500	503	Poly-Leu.
FT	CARBOHYD	54	54	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	132	132	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	138	138	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	244	244	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	416	416	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	431	431	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	464	464	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	479	479	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	497	497	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	506	506	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	626	626	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	723	724	RL -> GP (in isoform 3). /FTid=VSP_015312.
FT	VARSPLIC	725	1359	Missing (in isoform 3). /FTid=VSP_015313.
FT	VARSPLIC	1046	1072	Missing (in isoform 2). /FTid=VSP_015314.
FT	VARIANT	55	55	V -> L (nonresponse to ezetimibe treatment). /FTid=VAR_023369.
FT	VARIANT	1233	1233	I -> N (nonresponse to ezetimibe treatment). /FTid=VAR_023370.
FT	SEQUENCE	1359 AA;	148698 MW;	3225D53D93B76886 CRC64;
SQ				

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY	1	MAEAGLRGWLWALLRLAQSEPYTHIQPGVCAFVDECGKPNELSGSLMTLSNVSCLSN	60
QY	1	MAEAGLRGWLWALLRLAQSEPYTHIQPGVCAFVDECGKPNELSGSLMTLSNVSCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTHIQPGVCAFVDECGKPNELSGSLMTLSNVSCLSN	60
QY	61	TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF	120
QY	121	VNLHCNTCSPNQSLFINVTRVAQLGAGOLPAVAYEAPYQHSFAPQSDSCSRVRPAA	180
Db	121	VNLHCNTCSPNQSLFINVTRVAQLGAGOLPAVAYEAPYQHSFAPQSDSCSRVRPAA	180
QY	181	ATLAVGTMCVYGSALCNQRLNFGDGTGNGLAPLIDITFHLEPQAVGSGIQLPNEG	240
Db	181	ATLAVGTMCVYGSALCNQRLNFGDGTGNGLAPLIDITFHLEPQAVGSGIQLPNEG	240
QY	241	ARCNEQSGDDVATCSQDCAACPAIARPAQALDSTFYLQGMFGSLVLIILCSVFAVVTI	300
Db	241	ARCNEQSGDDVATCSQDCAACPAIARPAQALDSTFYLQGMFGSLVLIILCSVFAVVTI	300
QY	301	LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSFTHLLGQFPQGWGTWVASWPLTILVLS	360

DR	EMBL; AY437867; AAR97888.1; -; mRNA.	Db	120	FVSLHCHNTCS	PDQSLFINVTRVVERGAGEPPAVVAAYEAFYQKSPFAEKAYESCQVRIPA	179
DR	RGD; I303135; Npc111.	Qy	180	AATLAVGTW	CGVYSALCNAORWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEG	239
DR	InterPro; IPR004765; NP_C type.	Db	180	AASLAVGSM	CGVYSALCNAORWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEG	239
DR	InterPro; IPR003392; Patched.	Qy	240	VARNESQGD	DDVATSCQDCACAPAIARPAQALDSTFYLGQMPGSLVLIILCSVFAVVT	299
DR	InterPro; IPR000731; SSD 5TM.	Db	240	IAPCNESQ	GDSDSVCSCQDCACAPVIPPPEALRPSFYMGMPGWLALIIIFTAVFVLLS	299
DR	Pfam; PF02460; Patched; 1.	Qy	300	ILLVGRVAP	ARDKSMVDPKKTSLSDKLSFSTHLLGQFFQCGWGTWASWPLTIVLS	359
DR	TIGRFamB; TIGR00917; 2A060601; 1.	Db	300	AVLRLRV	VSNRNKNAEGSQEAPKLPKHKLSPHTILGRFQNGWGRVASWPLTIVLS	359
DR	PROSITE; PS0156; SSD; 1.	Qy	360	VIPVALAAG	IVTETLTDPELWSAPNSQARSEKAFHDQHPFPFRTNOVILTAPNRSS	419
DR	Cholesterol metabolism; Glycoprotein; Lipid metabolism; Signal;	Db	360	FIVVIALA	AGLTIELTDPVELWSAPKQARKEKSFHDEHPFPFRTNOJFVTARNRSS	419
DR	Steroid metabolism; Transmembrane.	Qy	420	YRYDSL	LGLPKNFGIILDLLELLELQERLRLQVWSPAQRNLSQDICVAPLNPDN	479
DR	SIGNAL	Db	420	YKDSL	LLGSKNFGIILDLLELLELQERLRLQVWSPAQRNLSQDICVAPLNPN	479
DR	CHAIN	Qy	480	TSYDCC	INSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTKDGTALA	539
DR	TOPO DOM	Db	480	TSLSDC	CVNSLLOVFQNNRTLLMTANQTLNGQTSYLVMDKDFYCANAPLTKDGTSLA	539
DR	TRANSMEM	Qy	540	LSCNADY	GAPVPELAIKGGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEE	599
DR	TRANSMEM	Db	540	LSCNADY	GAPVFPPLAVGGYQGTDYSEAEALIIIFSLNNYPADDPRAQAKLWEEAFLEE	599
DR	TRANSMEM	Qy	600	MRAFQRM	MAGMFQVTFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISALGYSYNSWR	659
DR	TRANSMEM	Db	600	MESQRT	SDKQVAFSAERSLEDEINRTTIQDLPEFAVSYIIIVELYISALGYSYNSR	659
DR	TRANSMEM	Qy	660	VMVDSKAT	LGGVAVVILGAVMAANGFYSYLGIRSSLYLOVVPFLVSLVACADNIFVFL	719
DR	TRANSMEM	Db	660	VAVESKAT	LGGLGVIVLGAAMGFYSYLGVPSSLYIIQVVPFLVAVGADNIFVFL	719
DR	TRANSMEM	Qy	720	EYQRLPR	PRGPREVHIGRALGRVAPSMMLCSLSEALICFFLGALTPMPAVTFALTSLA	779
DR	TRANSMEM	Db	720	EYQRLPR	MPGQREAHIGRTLGSVAPSMMLCSLSEALICFFLGALTPMPAVTFALTSLA	779
DR	TRANSMEM	Qy	780	VILDFLO	MSAFVALLSLDSKQBRASRLDVCCCVKQBELPPGQEGELLLGFFQKAYAPF	839
DR	TRANSMEM	Db	780	IILDFLO	MTAFVALLSLDSKQBRASRPDLVCCFSTRKLPPLPKKEGELLLRFRKRIYAPF	839
DR	TRANSMEM	Qy	840	LLHWITR	GVVLLFLALPGVSLYSNCHISVGLDQELALPKDSYLLDYFLNRYPEVGP	899
DR	TRANSMEM	Db	840	LLHFRF	REVMMLFLTLFGANLYMCNINVGDLQELALPKDSYLLDYFLNRYLEVGP	899
DR	TRANSMEM	Qy	900	VYFVTTL	GVNFSSEAGMNAICSSAGCNMFSTQKIQYATEPPEQSYLAIPASSWYDDFD	959
DR	TRANSMEM	Db	900	VYFVTTS	GVNFSSEAGMNAICSSAGCKFSLTQKIQYASEPDDQSYVAIAASSWYDDFD	959
DR	TRANSMEM	Qy	960	WLTP	-SSCCRLYISQPNKDKFCPTVNSLNCNCKMWSITMGSVRPSVQFHKYLPWFLND	1018
DR	TRANSMEM	Db	960	WLTPSS	CCRLYIRGPHKDEFCPTDTSFNCLKNCNRTLGPVREPTAQFHKYLPWFLND	1019
DR	TRANSMEM	Qy	1019	RNIKCPK	GGGLAAYSTVNLSTDGOVLASRFMAVHKPLKNSQDYTEALRAARELANITA	1078
DR	TRANSMEM	Db	1020	PNENICP	KGGLAAYSTVNLSSDGVVIAQFMAVHKPLKNSQDYTEALRASLANITA	1079
DR	TRANSMEM	Qy	1079	DURKVP	GTDPAPFVPPYITTVNFYEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRG	1138
DR	TRANSMEM	Db	1080	DURKVP	GTDPNEVEPYITISNVFYQYLTVPLEGIFTALCFVPTFVVCYLLGLDMCSG	1139
DR	TRANSMEM	Qy	1139	LNLLS	IVMILVDTVGFMALWDISYNAVSLINIVSAGVMSVEFVSHITSPATKPTWL	1198
DR	TRANSMEM	Db	1140	ILNLLS	IIIMILVDTIGLMAVNGISYNAVSLINIVTAVGMSVEFVSHITSPAVSTKPTL	1199
DR	TRANSMEM	Qy	1199	ERAKAT	ISWGSAGVAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLIITLLGLHGLVFL	1258
DR	TRANSMEM	Db	1200	ERAKOAT	VFMGSAGVAGVAMTNFPGLIILGFAQAQLIQIFFRLNLLIITLLGLHGLVFL	1259

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
Qy 1 MAEAGLRGWLWALLLRLAQSEPTTTHQPGYCAFYDECKNPGLSGMLTSLNVSCLSN 60
Db 1 MARAALW-GWLWALLLSAAQGLYTPKHEAGVCTFYEECKNPGLSGMLTSLNVSCLSN 59
Qy 61 TPARKITGDHLLILOKICPRLYTGPNQ-ACCSNKLVSLEASITTKALLTTCPCACSDN 119
Db 60 TPARKITGDHLLILOKICPRLYTGPNQ-ACCSNKLVSLEASITTKALLTTCPCACSDN 119
Qy 120 FVNLHCHNTCSPNQSLFINVTRVVAQLGAGQLPAVVAAYEAFYQKSPFAEKAYESCQVRIPA 179

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
QY	4	AGLGLWLLALLRLAQSEPTTTHOPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA	63
DB	3	AAWQGLWLLALLNSAQELTPTTHKAGCFYFECGKNPELSGGLTSLNSISCLNTPA	62
QY	64	RKITGDHLLILQKICPRLYTGN-TQACCSAKQLVSLASISYKALLTRCPACSDNFVN	122
DB	63	RHVTGDHLLALLQVCPRLYNGPNDTYACCSKQLVSLDSSLISITKALLTRCPACSENFVS	122
QY	123	LHCHTCSNQSLFNVTRVAQLGAGQLPAVVAEAFYOHSPAQSDYDSRVRVPAAT	182
DB	123	IHCNTCSPDOSLFNVTRVQRPDQLPAVVAEAFYQRFPAKAYECSRVRIPAAAS	182
QY	183	LAVGTMCGVYGALCNAORWLNFOQDTGNGLAPLDTIHFLLPEGQAVGSGIQLNPGVAR	242
DB	183	LAVGMCGVYGALCNAORWLNFOQDTGNGLAPLDTIHFLLPEGQALADGMKPLDGKITP	242
QY	243	CNESQGDVATCSCQDCAAACPAIARPOALDSTFFYLGOMPSLVLIILCSVFAVVTILL	302
DB	243	CNESQGEDSAACSCQDCAAACSPVIPPFPALRPSFTYMGPMGWLALIIFTAVFVLLSVVL	302
QY	303	VGRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQFQGGTGWASWPLTILVLSVIP	362
DB	303	VYLRVASNKNKNTAGSQAPNLDPKRKPSPTVLGRFFESGTRVAVSWPLTVLALSIV	362
QY	363	VVALAAGLVFTLTDPVVELSNAPSQAARSEKAFHDQHPFPFRFRTNVILTAHNRSSVRY	422
DB	363	VIALSVGLTFTLTDPVVELSNAPSQAARSEKAFHDQHPFPFRFRTNVITAKNRSSYKY	422
QY	423	DSLLGKPNFSGILDLDLLELELERLRLHQVSPSAQRNLSQDICYAPLNPDNTSL	482
DB	423	DSLLGKPNFSGILSDLLQELLELERLRLHQVSPSAQRNLSQDICYAPLNPHNTSL	482
QY	483	YDCINSLLQYFONNRTLLLTANOTLMGOTSQVMDKHFLYCANAPLTFKDGTLALSC	542
DB	483	TDCCVNSLLQYFONNHTLLLTANQTLNGQTSLVMDKOHFLYCANAPLTYKDGTLALSC	542
QY	543	MADYGAPVPFPLAIGGYKGDKYSEAEALIMTFSLNPNYPAGDPRLAQAQKWEAFLEMPA	602
DB	543	IADYGAPVPFPLAVGGVQGTDYSEAEALITFSINNYPADPRMAHAKLWEEAFLKEMQS	602
QY	603	FORRMAGMFOVTTAERSLEDEINRTAEDLPFTATSYVILYLSLALGSSYSSRVMV	662
DB	603	FORSTADKFQIAFSAERSLEDEINRTIQDLPVFAISYLVILYLSLALGSSYSSRVMV	662
QY	663	DSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLEYQ	722
DB	663	DSKATLGLGVAVVLGAVVAAMGFYSLGVPSLVILQVVPFLVLAAGADNIFIFVLEYQ	722
QY	723	RLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAICTFLGALTMPAVRTFALTSGLAVAL	782
DB	723	RLPRMPGEQREAHIGRTLGSVAPSMLLCSLSEAICTFLGALTSMAPVRTFALTSGLAIFP	782
QY	783	DFLLQNSAFVALLSLDSKREASRLDVCCCKQELPPGQGGGLLGFQKAYAPFLH	842
DB	783	DFLLQMTAFVALLSLDSKREASRPDVCCFSERNLPPPKQKGLLCLCFKRYITFLH	842
QY	843	WITRGVLLALFLALFGVLSYSMCHISVGLDQELAPKDSYLLDYFLNRYFVGPAPVP	902
DB	843	RFIRPVVLLFLVLFGANLYLMCNISVGLDQELAPKDSYLYDYFLNRYLSEVGPVPV	902
QY	903	VTTLYGNFSSPAGNNAICSSAGCNPFSTQKIYATEFPFQSYLAIPASSWDDFDLWLT	962
DB	903	DTTSGYNFSTBAGNNAICSSAGCSFSLTQKIYASEFPNQSYVAIAASWDDFDLWLT	962
QY	963	P-SSCCRIYISGPNKDKFCPSVTNLSCLNKMISITMGVSVRPSVEQPHKTLPLWFLNDRPN	1021
DB	963	PSSCCRIYTRGPHDKDFCPSDTSTFNCLNKMRTILGPVRPTTEQPHKYLPLWFLNDRPN	1022
QY	1022	IKCPKGGLAAYSTSVNLTSQGVLAASRPMAHYKPLKNSQDYTEALRAARELANITADLR	1081
DB	1023	IRCPKGGLAAYVRTSVNLSSDQIIIASQPMAYHKPLRNSQDFTALRASRLAANITAEALR	1082

QY	1082	KUPGTDPAPEVPPYTTINVFYQYLTLPEGLPMLSLCLVPTFAVSCLLGLDRLSGLLN	1141
DB	1083	KVEGTDNPEVPPYTTISNVFYQYLLTLVPEGLFTLALCFVPTFVVCYLLGLDRLSGLLN	1142
QY	1142	LLSIVMLVDVTGCFMALWDISYNAVSLINLVSAGMVSVEFVSHITSFSAISTKPTWLER	1201
DB	1143	LLSIIIMLVDTIGLMAVWGISYNAVSLINLVTAVGMSVEFVSHITSFSAVSTKPTRLER	1202
QY	1202	KEATISMGSAVFAVAVMTNLPGILVTLGAKAQILQIFPRLNLLITLLGLLHGLVLPVI	1261
DB	1203	KDATICWGSAAVFAVAVMTNPPGILILGFAQAQILQIFPRLNLLITLLGLLHGLVLPVV	1262
QY	1262	LSVVGDPNVPALALEOKRAEEAAVAVWVASCNPNHPSKRVSTADNIIYVNHSPGS-IGAGA	1320
DB	1263	LSVLGDPNVALVLEEKLEATEA-AMVSEPCQPPPADANTSDYVNYGNPFIFEINA	1321
QY	1321	ISNFPNNGRQF 1332	
DB	1322	ASSSLPKSDQKF 1333	
RESULT 4			
ID	Q4T749	TETNG PRELIMINARY; PRT; 1132 AA.	
AC	Q4T749		
DT	13-SEP-2005	(TREMBlrel. 31, Created)	
DT	13-SEP-2005	(TREMBlrel. 31, Last sequence update)	
DE	13-SEP-2005	(TREMBlrel. 31, Last annotation update)	
DE	Chromosome	undetermined SCAF8317, whole genome shotgun sequence.	
DE	(Fragment).		
GN	ORFNames=GSTENG0005930001;		
OS	Tetraodon nigroviridis (Green puffer).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;		
OC	Tetraodontidae; Tetraodontidae; Tetraodon.		
OX	NCBI_TaxID=99883;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Jailion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,		
RA	Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,		
RA	Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,		
RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,		
RA	Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,		
RA	Bienmont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,		
RA	Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,		
RA	Parra G., Lardier G., Chapple C., McKernan K.J., Mesirov J.,		
RA	Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,		
RA	Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,		
RA	Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,		
RA	Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,		
RT	"Genome duplication in the teleost fish Tetraodon nigroviridis reveals		
RT	the early vertebrate proto-karyotype.";		
RL	Nature 431:946-957(2004).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RG	Genoscope; Whitehead Institute Centre for Genome Research;		
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.		
CC	-!- CAUTION: The sequence shown here is derived from an		
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is		
CC	preliminary data.		
DR	EMBL; CAAB01008317; CAF91283.1; -; Genomic_DNA.		
FT	NON TER 1		
SQ	SEQUENCE 1132 AA; 125923 MW; 4283CF4CA0BDE97 CRC64;		

Query Match 46.0%; Score 3179; DB 2; Length 1132;			
Best Local Similarity 52.0%; Pred. No. 4.3e-216; Gaps 11;			
Matches 641; Conservative 167; Mismatches 264; Indels 160;			
QY	87	TOACCSAKQLVSLASLSITKALLTRCPACSNFVNLHCHNTCSNPQSLFINVTRVAQL-	145
DB	1	TYACCSINQLSLEASLALS KAVLRCPSCADNFAHFHCITTCSPQQTIKTKVTNVT	60

Db 4 RGPAGLILLLLCPVQFSQ-----SCWVYGECC---IASGD-----KRYNCRYSQPPKPL 51
Qy 67 TGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCH 126
Db 52 PEDGYDLVQELCPGFFG-NVSLCCDVQQLRTLDKDLPLQLFLSRCPSCFFNLNLFCE 110
Qy 127 NTCSPNOSLFINVTR-----VAQLGAGQLPAAVYEAFFYQHSFABQSDSCSRVRPAAAT 182
Db 111 LTCSPROQFLNVTATEDYDVPVNTQNTKTNVKELEYVYGETPANAMYNACRDVEAPSSNE 170
Qy 183 LAVGTMCGVYGSAICNAQRWLNFOGDTGNGLAPLDIT--FHLLEPGQAVGSGIQPLNEG 240
Db 171 KALGILLCGREQA-CNATNWIYMFNKGQAPFTITPISDLPT-----HGMPEPMNAT 224
Qy 241 ARCNSQGDVATCSQDCAASC-----PAIARQALDSTFYLGQMPGSLVLIIL 291
Db 225 KGCDESDEVGTGPCQDCSIVCGPKQPPPPVPMRILGLDAMVIMWSSYMAFLIYVP 284
Qy 292 CSVFAVTVILLGVFAVAPARDKSKMDPKGTSLSKLSFSTHTL----- 336
Db 285 GAFFAVWCY-----RKRYFVSEYTPIDGNIAFVNSSDKGQAFCCDPLGAAF 331
Qy 337 ---LGQFFQGGTGWASWPLTILVLSVIPVALAAGLVFTLTTPDVELMSPNSQARSE 393
Db 332 ERGLRLFAQWACFVRHPGCVFFSLAFIVACSSGLVFIIRVTTDPDVLMSAPGQARRE 391
Qy 394 KAFDHQHFPPFRMTQVILTAPNRSYRYDSLLGLPK-NFGSILDLILLLELLEQLERL 452
Db 392 KEYFDTHGPFPRMEQLIRATNNGSHLYHPYPAGADVPPGPPLSRDLHQVLDLQTAIE 451
Qy 453 HLQWSPAQRNISLQDI CYAPLNPDTNLSVDCINSILQVFNQNRITLLLTANQTLMGQ 512
Db 452 N--ITASYNNETVTLQDLCAPLSYKN---NCTILSVLNYFQNSHSLV----- 495
Qy 513 TSQV-----DWKDHFLYCANAPLTFKDGTALASCWADYGAPVPPFLAIGVKGKDY 564
Db 496 DHQGDFFVADVHTHFLYCVRAPASLNDASLLHDCPLGTGFGVPPFWLVGLGVDYQNY 555
Qy 565 SEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLERAFQRRMAGMFOVTTAERSLDE 624
Db 556 NNATALVITPPNNYNDTEKLQRAQWESFNFVKNYK---PNLTISFWAERSIDE 612
Qy 625 INRTAEDLPATSYIVIFLYISALGSSWSRWVDSKATGLGVAVVVLGVAAMAA 684
Db 613 LNRESNLDLITLISYAIMFLYISIALGHKSCSRLVDSKISLGIAGILVLSVACSL 672
Qy 685 GFESVIGTRSSIVILQVPELVLSGADNIEFVLEYQRLRRRPGEPREVIHGRALGVA 744
Db 673 GIFSYIGVPLTLVIEVIFPLVLAAGVDNIIFILVQTYQYQDERLOQETLDQQLGRVLGVA 732
Qy 745 PSMILICSLSEACFPPLGALTMPAVRTPALTSGLAVILDFLLQMSAFVALLSLDSKQEA 804
Db 733 PSMFLSSSEVAVPLGLSVVPAVHTSLPAGMAVLIDPLLOITCFVSLGLDLIKRQEK 792
Qy 805 SRLDVCCCKPQELPPPGQ---EGLLGFQKAYAPFLHWTIRGVVLLFLFALFGVS 860
Db 793 NRLDVCCVQGAEE---DGAGVQAESCLFRFFKNSYAPLLKDWMRPIVAVFVGVLSFS 849
Qy 861 LYSMCHISVIGLDOELALPKDSVLLDYFLFNRYFVGVGAPVYFVTLGNPFSSEAGMAIC 920
Db 850 IAVLNKVBIGLDQSLSPDDSVYMDYFQSLGRYLHAGPPVYFVVEBGNHYTSLKGQNMVC 909
Qy 921 SSAGCNFSFTQKIYATPEPQSYALIPASWVDDFDLWLTTP-SSCCRLYISGNPKDKF 979
Db 910 GGLGNNDSLIVQIPTAQLDNYTIFGAPASWIDDDYEDNFKPQSSCCRVNS---TDQF 966
Qy 980 CPSTVNSLNCNKMCSITM-GSVRPSVBQPHKYLPMFLNDRPNIKCPKGGLAAYSTS 1038
Db 967 CNASVVDPTCIR-CRPLTSEGKQRFQGEDFMFLPMFLSDNPNPCKGGHAAYSVA 1025
Qy 1039 TSDGO-VLASRPMAYHKPLKNSQDYTEALRAARELANITADLRKVPDTPDAFEVFPYTI 1097
Db 1026 LKNGSGVGATYFMTYHTVLOASADPIDAMQKARLIASNIT---RTWGLEASSYRVPYYSV 1082

Qy 1098 TNVFEQYLTITLPEGLFMLSCLVPTFAVSCLLGLDRLSGLNLLSLVIMLVDTVGPMA 1157
Db 1083 FVVFYEQYLTVIDDTIFNLGVSGLAIFLAVTVLWGCCELWATVIMCVTIAMILVNMFGVM 1142
Qy 1158 LWDISYNAVSLINLVSAGMSVEFVSHITRSPALSTKPTWLERAKEATISWGSAPVAGA 1217
Db 1143 LMGISLNAVSLINLVMSCGISVBFCSHITRAFTLSTKGSRDVRAEALAHMGSVFSGIT 1202
Qy 1218 MTNLPGILVLGLAKAQLIQIFFFRLNLLITLLGLLHGLVFLPVLISVYVGPDPVNPALAL 1275
Db 1203 LTKFGGIVLAFAPKASQIFQIFFRMYLAIVLLGATHGLIFLPLVLSYIGPSINRAKSL 1260
RESULT 7
Q8MKD8_FELCA
ID Q8MKD8_FELCA PRELIMINARY; PRT; 1276 AA.
AC Q8MKD8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Niemann-Pick C1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
OC Felis.
OX NCBI_TaxID=9685;
RN (1)
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=22695580; PubMed=12809639; DOI=10.1016/S1096-7192(03)00074-X;
RA Somers K.D., Royals M.A., Carstea E.D., Rafi M.A., Wenger D.A.,
RA Thrall M.A.;
RT "Mutation analysis of feline Niemann-Pick C1 disease.";
RL Mol. Genet. Metab. 79:99-103(2003).
DR EMBL; AF503634; AA027451.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; P:hedgehog receptor activity; IEA.
DR GO; GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP_C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
DR SEQUENCE 1276 AA; 141710 MW; C2DD3496472B71EP CRC64;
Query Match 34.8%; Score 2405; DB 2; Length 1276;
Best Local Similarity 40.7%; Pred. No. 3.8e-161;
Matches 535; Conservative 232; Mismatches 461; Indels 86; Gaps 29;
Qy 7 RGMILLALLLLAQSEPTTTHQPGYCAFYDECGKNPBLSGSLMTLSNVSCLSNTPARKI 66
Db 4 RGPALGILLLLC---PAQVLAQS---CIWYGECC---IASGD-----KRYNCRYSQPPKPL 51
Qy 67 TGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCH 126
Db 52 PEDGYDLVQELCPGFFG-DNVSLCCDVQQLRTLDKDLPLQLFLSRCPSCFFNLNLFCE 110
Qy 127 NTCSPNOSLFINVTR-----VAQLGAGQLPAAVYEAFFYQHSFABQSDSCSRVRPAAAT 182
Db 111 LTCSPROQFLNVTATEDYDVPVNTQNTKTNVKELEYVYGESFANAMYNACRDVEAPSSND 170
Qy 183 LAVGTMCGVYGSAICNAQRWLNFOGDTGNGLAPLDIT--FHLLEPGQAVGSGIQPLNEG 240
Db 171 KALGILLCGKDAE-CNATNWIYMFNKGQAPFTITPISDLPT-----HGMPEPMNAT 224
Qy 241 ARCNSQGDVATCSQDCAASC-----PAIARQALDSTFYLGQMPGSLVLIIL 291
Db 225 KGCDESDEVGTGPCQDCSIVCGPKQPPPPVPMRILGLDAMVIMWITMAFLVLPFF 284
Qy 292 CSVFA-----VVTILLGVFA---PARDKSKMDPKGTSLSKLSFSTHTLLGQ 339

Db	285	CAFFALMCKYRKRYFVSEYTPIDSNIAFVSNDRGE-----ASCCDALGAAPGECCLR	337
Qy	340	FFQGGTWTWASWPLTILVLSVIPVVALAAGLVFTELTPVELWSAPNSQARSEKAFHDQ	399
Db	338	LPSQWGSFCVNPGLPIFFSLAFTACSSGLVFRVTNPVDLWSAPSSQARLEKEYFDT	397
Qy	400	HFGPFRTNQVILTAPNRSSRYDSSLGPK-NFSGILDLDLLELELERLHLQVMS	458
Db	398	HFGPFRTQVILTAPNRSSRYDSSLGPK-NFSGILDLDLLELELERLHLQVMS	455
Qy	459	PEAQRNISLQDICVAPLNDWTSYDCCINSLQYFQNNRTLLLTANQTLMGQTSOV--	516
Db	456	SYNNTVTLQDICVAPLNDWTSYDCCINSLQYFQNNRTLLLTANQTLMGQTSOV--	507
Qy	517	DWKDHFYCANAPLTFKDGOTALALSCMADYAGAPVFPFLAIGYKGDYSEAEALIMTFSL	576
Db	508	DYTHLLCYVRAPASLNTSLHDFCLGTGPGVFPWLVLGGYDDQNNYNATALVITFPV	567
Qy	577	NNYPAGDPRLAQAKLEAFLEEMAFQRMAGMFQVTTAERSLEDBINRTTAEPLIF	636
Db	568	NNYNDTERLOKHWEKEFINFVNKNYKN---PNLTISFTTTERSTEDLNRESNGDFTV	624
Qy	637	ATSYVIVIFYISLALGSYSSSRVMVDSKATLGLGVAVLGAVMAAGFFSYLGRSSL	696
Db	625	IISYAIMFLYISALGHKSCSRLVDSKISLIGIAGILIVSSVACSLGIESYVGIPLTL	684
Qy	697	VILQVVPFLVLSVAGADNIFIFVLEYQRLPRRPEGEPRVHIGRALGRVAPSMLLCSLSPAI	756
Db	685	IIVIEVIFVLAVGVNDVNFILVQYQDERLHGETLDQQLGRVLGEVAPSMFLSFSBAV	744
Qy	757	CFPLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSDSKQBASRLDVCCKVPQ	816
Db	745	APFALGALSMPAVHTFSLFAGMAVLIDFLQITCFVSLGLDIXKQEKRLDVLCCVRGS	804
Qy	817	ELPPPPQ-CEGILLGFFQKAYAPPELLHMTTRGVLLLLFLALFGVSLYSCHISVGLDQEL	875
Db	805	EDGTSVQASCESLFRFLFKHSYSPLLLKQWRPFIATFVGVLSFVAVLNKVEIGLDQSL	864
Qy	876	ALPKDYSYLDYFLFLNRYFEVGPVYFVTTGLVNFSSSEAGMNAICSSAGCANNFSTQIKI	935
Db	865	SWPDDSYVMDYFKSL-KYLHAGPPVYFVLEEGHDYTSLKQNNVCGMGCCNDSLVQQTIF	923
Qy	936	YATEFPEQSYALPAGSWDDFDWLTTP-SSCRLYISGPNKDKCPCSTVNSLNCIKNCM	994
Db	924	NAAQLDSTYRIGFAPSSWIDDYDWDYKQSSCCRVNYS---TDRFCNASVDPACIR-CR	979
Qy	995	SITM-GSVRPSVQFHKYLPWFLNDRPNKCPKGLAAYSTSVN-LTSDGQVLASRFMAY	1052
Db	980	PLTQEGKQRQGGDFMRFLPMFLSDNPNPKCGKGHAAYSSAVNLLGNDTGVGATYFMTY	1039
Qy	1053	HKPLKNSQDYTEALRAARELANITADLRKVPQTDPA-FEVFPYTTTNVFPYQYLILPE	1111
Db	1040	HTVLTQTSADFTDAMRKANLIASNIT---KTMGLEGSNRYVPYFVYFYEQYLITIID	1095
Qy	1112	GLFMLSCLVPTTAVSCLLGLDLRSLGNLLSVMLVLDVTCGFMALWDISYNAVSLNL	1171
Db	1096	TIFNLSVSLGAIFLVTVILGCDLWSAVIMCITIAMILVMFQVMWMLGSLNANSLVNL	1155
Qy	1172	VSAGVMSVZFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVILGLAK	1231
Db	1156	VMSCGISVFCSHITRAFTVSMKGSRAQRAEALAHMSSVFGSITLTKPGGIWVLAFAK	1215
Qy	1232	AQLTQIFFFRLNLLITLGLLHGLVFLPVLISVGPVDPVNPALAI---EQKRAE	1282
Db	1216	SOIFQIFVFRMYLANVLLGATHGLIFLVLVLSYIGPSINKAKSLATQEQYKGTG	1269
RESULT 8			
Q9GK52 CANFA PRELIMINARY; PR; 1276 AA.			
AC Q9GK52;			
DT 01-MAR-2001 (Tremblrel. 16, Created)			

Db 583 WEKEFNFKVKNYEN---PNLTISFTTTERSDIELNRESNGDVFVTLISYAVMFLYSIAL 639
QY 652 GSYSRSRVWDSKATLGGVAVVLGAVMAAGRPYSYLGIRSSLVILQVWVPLVLSVCA 711
Db 640 GHKSCSRFLVDSKISLGIAGILVLSWCSLGIYSFGIPDLTILVIEVIFLVLAVG 699
QY 712 DNIFIVLBYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSAICFFLGAITPMPAVRT 771
Db 700 DNIFILVQYQDERLQGETLEQQLGRVLGEVAPSMFLSSFEAVAFPLGALSQMPAVHT 759
QY 772 FALTSLAVILDFLOMSAFVALLSDSKQERASRLDVCCCKPQBLPPGQG----EGL 827
Db 760 FSLFAGMAYLIDFLQITCFVSLGLDLKRQENRDLVLCCLTGSB---GGTGIQASESC 816
QY 828 LIGCFKOKAVAPFLHWHITGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYF 887
Db 817 LRFFFKNSVPELKKDWMPVIANVFGLISFIAVLNKVEIGLDQSLMSDDSYNMDYF 876
QY 888 LFLNRYFEVGAPYFVFTTLGYNFSSEAGNNAICSSAGCNFFSTQKIQYATFPPEQSYLA 947
Db 877 KSL-KYLHAGPPVYFVLEBHDYTSLEGQNMVCGMGCCNDSLVQIIPSAQOLDNYTHIG 935
QY 948 IPASSWDDFIDLTP-SSCRLYISGPNKDFPCSTVNSLCLNKNMSITM-GSVRPSV 1005
Db 936 FAPSSWIDDYDFWVKPQSSCCRVTNS---TDQFCNASVVDPAQVR-CRPLTQGRKRPQG 991
QY 1006 EOPHYKLPWFLDRPNKCPKGLAAYSTSVNLT-SDGOVLASRMAVYHKLKNSQDYTE 1064
Db 992 EDPMPFLPFLSDNPNPKCGKGAAGVAVNLVNDTSVGGATYFMTYHVLVJQTSADFID 1051
QY 1065 ALRAARELAANTADILRKVPGTDPAPFVPPYITITNFFYEQYLTILPEGLFMLSCLVPTFF 1124
Db 1052 AMRKAILIASNITKM-SLKGS-VRVPYSVYFVYEQYLTIDDTIFNLVSLSGATP 1108
QY 1125 AVSCILLGLDLSGLNLNLISVILVDTVGFMAIDISVAVSLNLINLAVSGMSVEFVSH 1184
Db 1109 LVTLLVLLGCELMSAVIMCVTAMILVMPGVMLWGLISLAVSLVNLVNSCGISVEFCSH 1168
QY 1185 ITRSPAISNPTWLRERAKATISMSGAVGAVAMTNLPGILVLGLAKAQLIIFPRLNL 1244
Db 1169 ITRAFTVSANGSRVEREALSHMSSVFSGLTKFGGIVVLAFAKSQIFQVFTGMYL 1228
QY 1245 LITLLGLHLGLFVLPSYVGPDPVNPALAL-EQKR 1279
Db 1229 ANVLGATHGLIFLPVLSYIGPSINKAKSLASQER 1264
RESULT 9
ID NPC1 HUMAN STANDARD; PRT; 1278 AA.
AC O15178; Q9P130;
DT 30-MAY-2000 (Rel. 39, Created)
DE 10-MAY-2005 (Rel. 47, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Niemann-Pick C1 protein precursor.
GN Name=NPC1;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, VARIANT ILE-642, AND VARIANTS NPC1.
RX MEDLINE=97362323; PubMed=9211849; DOI=10.1126/science.277.5323.228;
RA Carstea E.D., Morris J.A., Coleman K.G., Loftus S.K., Zhang D.,
RA Cummings C., Gu J., Rosenfeld M.A., Pavan W.J., Krizman D.B.,
RA Nagle J., Polymeropoulos M.H., Sturley S.L., Ioannou Y.A.,
RA Higgins M.E., Comly M., Cooney A., Brown A., Kaneski C.R.,
RA Blanchette-Mackie E.J., Dwyer N.K., Neufeld E.B., Chang T.-Y.,
RA Liscum L., Strauss J.F. III, Ohno K., Zeigler M., Carmi R., Sokol J.,
RA Markie D., O'Neill R.R., van Diggelen O.P., Ellender M.,
RA Patterson M.C., Brady R.O., Vanier M.T., Pentchev P.G., Tagle D.A.;

RT "Niemann-Pick C1 disease gene: homology to mediators of cholesterol
RT homeostasis.";
RL Science 277:228-231 (1997).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND VARIANTS.
RX MEDLINE=99355599; PubMed=10425213; DOI=10.1006/bbr.1999.1070;
RA Morris J.A., Zhang D., Coleman K.G., Nagle J., Pentchev P.G.,
RA Carstea E.D.;
RT "The genomic organization and polymorphism analysis of the human
RT Niemann-Pick C1 gene.";
RL Biochem. Biophys. Res. Commun. 261:493-498 (1999).
RN [3]
RP NUCLEOTIDE SEQUENCE, AND VARIANT ILE-642.
RX MEDLINE=21623216; PubMed=11754101; DOI=10.1002/humu.10016;
RA Bauer P., Knoblich R., Bauer C., Finckh U., Hufen A., Kropp J.,
RA Braun S., Kustermann-Kuhn B., Schmidt D., Harzer K., Rolfs A.;
RT "NPC1: complete genomic sequence, mutation analysis, and
RT characterization of haplotypes.";
RL Hum. Mutat. 19:30-38 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT ILE-642.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=1247932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzywicki M.I., Skaleka U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=99128318; PubMed=9927649; DOI=10.1073/pnas.96.3.805;
RA Watari H., Blanchette-Mackie E.J., Dwyer N.K., Glick J.M., Patel S.,
RA Neufeld E.B., Brady R.O., Pentchev P.G., Strauss J.F. III;
RT "Niemann-Pick C1 protein: obligatory roles for N-terminal domains and
RT lysosomal targeting in cholesterol mobilization.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:805-810 (1999).
RN [6]
RP VARIANT NPD TRP-992.
RX MEDLINE=98299797; PubMed=9634529;
RA Greer W.L., Riddell D.C., Gillan T.L., Girouard G.S., Sparrow S.M.,
RA Byers D.M., Dobson M.J., Neumann P.E.;
RT "The Nova Scotia (type D) form of Niemann-Pick disease is caused by a
RT G3097-->T transversion in NPC1.";
RL Am. J. Hum. Genet. 63:52-54 (1998).
RN [7]
RP VARIANTS NPC1 GLN-934; LEU-940; ASN-948; TRP-992; ALA-1007;
RP THR-1061 AND VAL-1213.
RX MEDLINE=99452586; PubMed=10521290;
RA Greer W.L., Dobson M.J., Girouard G.S., Byers D.M., Riddell D.C.,
RA Neumann P.E.;
RT "Mutations in NPC1 highlight a conserved NPC1-specific cysteine-rich
RT domain.";
RL Am. J. Hum. Genet. 65:1252-1260 (1999).
RN [8]
RP VARIANT NPC1 THR-1061.
RX MEDLINE=99452593; PubMed=10521297;
RA Millat G., Marçais C., Rafi M.A., Yamamoto T., Morris J.A.,
RA Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;

RT "Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant
RT allele in patients of Western European descent and correlates with a
RT classic juvenile phenotype.";
RL Am. J. Hum. Genet. 65:1321-1329(1999).
RN [9]
RX VARIANTS NPC1, AND VARIANTS ARG-215; VAL-858 AND GLN-1266.
RP MEDLINE=99408226; PubMed=10480349; DOI=10.1007/s004390051057; Zhang H.,
RX Yamamoto T., Namba E., Ninomiya H., Higaki K., Taniguchi M., Zhang H.,
RA Akaboshi S., Watanabe Y., Takeshima T., Inui K., Okada S., Tanaka A.,
RA Sakuragawa N., Millat G., Vanier M.T., Morris J.A., Pentchev P.G.,
RA Ohno K.;
RT "NPC1 gene mutations in Japanese patients with Niemann-Pick disease
RT type C.";
RL Hum. Genet. 105:10-16(1999).
RN [10]
RX VARIANTS NPC1 GLN-958 AND ALA-1007.
RP MEDLINE=21313111; PubMed=11349231;
RX Sun X., Marks D.L., Park W.D., Wheatley C.L., Puri V., O'Brien J.F.,
RA Kratt D.L., Lundquist P.A., Patterson M.C., Pagano R.E., Snow K.;
RT "Niemann-Pick C variant detection by altered sphingolipid trafficking
RT and correlation with mutations within a specific domain of NPC1.";
RL Am. J. Hum. Genet. 68:1361-1372(2001).
RN [11]
RX VARIANTS NPC1 ALA-378; MET-950; ARG-992 AND THR-1061.
RP MEDLINE=21313105; PubMed=11333381;
RX Millat G., Marais C., Tomasetto C., Chikh K., Fensom A.H., Harzer K.,
RA Wenger D.A., Ohno K., Vanier M.T.;
RT "Niemann-Pick C1 disease: correlations between NPC1 mutations, levels
RT of NPC1 protein, and phenotypes emphasize the functional significance
RT of the putative sterol-sensing domain and of the cysteine-rich luminal
RT loop.";
RL Am. J. Hum. Genet. 68:1373-1385(2001).
RN [12]
RX VARIANTS NPC1 TYR-177; CYS-978 AND VAL-1035.
RP MEDLINE=21372069; PubMed=11479732; DOI=10.1007/s004390100531;
RX Ribeiro I., Marcao A., Amaral O., Sa Miranda M.C., Vanier M.T.,
RA Millat G.;
RT "Niemann-Pick type C disease: NPC1 mutations associated with severe
RT and mild cellular cholesterol trafficking alterations";
RL Hum. Genet. 109:24-32(2001).
RN [13]
RX FUNCTION: Involved in the intracellular trafficking of
CC cholesterol. May play a role in vesicular trafficking in glia, a
CC process that may be crucial for maintaining the structural and
CC functional integrity of nerve terminals.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Found in late
CC endosomes and lysosomes.
CC -1- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain
CC containing a di-leucine motif necessary for lysosomal targeting
CC are critical for mobilization of cholesterol from lysosomes.
CC -1- PTM: Glycosylated.
CC -1- DISEASE: Defects in NPC1 are the cause of Niemann-Pick disease
CC type C1 (NPC1) [MIM:257220]. NPC1 is an autosomal recessive lipid
CC storage disorder, which affects particularly the brain, liver and
CC spleen, and which is characterized by lysosomal accumulation of
CC low density lipoprotein derived cholesterol. Clinical features
CC include variable hepatosplenomegaly and severe progressive
CC neurological dysfunction such as ataxia, dystonia and dementia.
CC The age of onset can vary from infancy to late adulthood.
CC -1- DISEASE: Defects in NPC1 are the cause of Niemann-Pick disease
CC type D (NPD) [MIM:257220]; also known as Niemann-Pick disease
CC without sphingomyelinase deficiency, or Nova Scotian type. Because
CC of evidence from biochemical changes, lack of complementation, and
CC linkage mapping to the same chromosome site, NPD and NPC1 are
CC considered to be allelic disorders.
CC -1- SIMILARITY: Belongs to the patched family.
CC -1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; AF002020; AAB63982.1; -; mRNA.
DR EMBL; AF157379; AAD48006.1; -; Genomic DNA.
DR EMBL; AF157365; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157366; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157367; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157368; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157369; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157370; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157371; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157372; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157373; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157374; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157375; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157376; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157377; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157378; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF123046; AAF28875.1; -; Genomic DNA.
DR EMBL; AF123045; AAF28875.1; JOINED; Genomic DNA.
DR EMBL; BC063302; AAF63302.1; -; mRNA.
DR Ensembl; ENSG00000141458; Homo sapiens.
DR HGNC; HGNC:7897; NPC1.
DR MIM; 607623; -.
DR MIM; 257220; -.
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0005764; C: lysosome; TAS.
DR GO; GO:0005624; C: membrane fraction; TAS.
DR GO; GO:0005478; F: intracellular transporter activity; TAS.
DR GO; GO:0015248; F: sterol transporter activity; TAS.
DR GO; GO:0004888; F: transmembrane receptor activity; TAS.
DR InterPro; IPR004765; NP_C_type.

Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 5.8e-161; Gaps 25;
Matches 522; Conservative 237; Mismatches 450; Indels 103;

QY 7 RGLWLLWALLRLAQSEPYTHIQPGYCAFYDECG-----KNPFGSLMTLSNVSLCS 59
DB 4 RGLALGULLLLCPAQVFSQ-----SCWYGECEIAYGDKRYNCEYS----- 46
QY 60 NTPARKITGDHLLILLOKICPLRYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDN 119
DB 47 --PKPLPKDGYDLVQELCPGFFG--NVSLCCDVRQLQTLKDNQLQPLQFLSRCPSCFN 103
QY 120 FVNLCHNTCPNOSLFINVTR---VAQLGAGQLPAVYAEAFYQHSFASQSDSCSRV 175
DB 104 LLNLECELTCPRQSQFLNVTATEDYDVPNTKTNVKELQYVYGQSFANAMYNACRDV 163
QY 176 RYVPAATLAVGTMCVYGSALCNQRWLNFGQDTGNGLAPLDT-----PHLLEPGA 228
DB 164 EAPSSNDKALGILLCGKDADA-CNATNWTYMFNKNQNGQAPFTITPVFSDFPVH----- 215
QY 229 VCSGQPLNEGVARCNESQGDVATCSQDCAACPAIARPO-----ALDSTFYLG 279
DB 216 ---GHEPMNATKCDSEVDVETAPCSQDCSIVCGPKPQPPPPAPWTLGLDAMYIM 272
QY 280 QMPGSLVLIILICSVFAVVTILL-----VGRVAPAPKSMQVKKGTSL 326
DB 273 WITYWAPLLVFGAPFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLQOFGQGTWVASWPLTILVLSVIVVALAAGIVFETELTTPDVELWSAP 386
DB 325 DPVSAFAFEGCLRLFLTRWGSFCVRNPGCVIFFSLVFTACSSGLVFRVVTTPVDLWSAP 384
QY 387 NQARSEKAFHDPHGFPRPTNOVITLAPNRSYRSDLSLLGPK-NFSGILDLDLLELL 445
DB 385 SSOARLEKBYFDHGFPRPTTEQLIRAPLTKHYQPYPSGADVPFGPLDIQLHQLV 444
QY 446 ELQERLRHIVQVMSPEAQRNLSQDICYAPLNPDTNLSYDCCINSLLQYQNNRLLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLDICLAPLSPYNT---NCTILSVLNYFQNSHSLVHKK 499
QY 506 NQTLMGQTSQVDWKDHFLYCANAPLTFKDTALALSCMADYGAIPVFPFLAIGYKGDYS 565

Db 500 GDDFF---VYADYHTHFLYCVAPASLNDTSLHDFCLGTGCGPVPFVPLVGLGDDYDQNTN 556
 QY EAEALIMTSLNNYPAGDPRLAQAKLWEBAFLMEARAFORMAGMPOVFTFAERSLEDEI 625
 Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTISFTAERSIEDBL 613
 QY 626 NRTAEDLPFATSYIVIFLYISLALGSSYSSRSMVMDSKATGLGGVAVIAGVMAAG 685
 Db 614 NRESDDVFTWISYAIMELYISLALGHMKSCRLLVDSKVLGIAGILVSSVACSIG 673
 QY 686 PFSYLGISSVLQVPLVSVGADNIFIVLEYQRLPRPGEPRVHIGRALGRVAP 745
 Db 674 VFSYIGLPLTVIEVIFPLVAVGVDNIFILVQAYQYORDERLQGETLDQOLGRVLEAVP 733
 QY 746 SMLLCSLSBAICFFGALTPMAVRTFALTSGLAVILDFLLQMSAPVALLSDSKQRAS 805
 Db 734 SMFLSSFSSTVAFFLALSVMPAVHTFSLFAGLAVFIDFLLQITCFVSLGLDIKRQKN 793
 QY 806 RLDDCCVKBQLPPPGQ--GEGLLGFPQKAVAPFLLHWTITRGVLLFLFALFGVSLYSM 864
 Db 794 RLDFCCVGAEDGTSVQASESCLFRFFKNSYSPILLKDMWRPIVIAIFVGVLFSFIAVL 853
 QY 865 CHISVGLDQELALPKDSYLLDFLPLNRYFEVGAUVYFTTLGYNPSSSEAGNAICSSAG 924
 Db 854 NKVDIGLQSLMPDDSDYMDVFKSISQYLHAGPPVYFVLEBHDYTSKQGNMVGCGMG 913
 QY 925 CNNFSTQIKQVATEPPEQSYLAIPASSWVDDPIDWLT--SSCCRLYISGPNKDKFCPT 983
 Db 914 CNDSLVQOIFNAQLDNTYRGAPSSMIDYDFWVKQSSCCRV---DNITDQFCNAS 970
 QY 984 VNSLNLKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYTSVN--LTS 1040
 Db 971 VDDPACVR--CRPLTBEGKQRPQGGDFMFLPFLSDNPNPKCGKGGHAYSSAVNLLGH 1029
 QY 1041 DQVLAARMYAKPLKNSQDTEALRAARELAANITADLRKVGTDPAFEVFPYITNV 1100
 Db 1030 GTRVGATYEMTYHTVLQTSADPIDALKKARLIASNT--ETMINGS--AYRVFPYSVFV 1086
 QY 1101 FFEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLNLLSIWMLIVDTVGMALWD 1160
 Db 1087 FFEQYLTIDDTIFNLGVSIGAIFLVTMVLGELMSAVIMCATIAMVLNMPGVWLMG 1146
 QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSPAITKPTWLERAKEATISMGSAVAGVAMTN 1220
 Db 1147 ISLNAVSLNLMSCGISVEFCSHITRAFTVSMKGSRVERAEBALAHMSSVFSGITLTK 1206
 QY 1221 LFGILVLGLAKQLQIFFRNLNLLITLLGLHLVFLPVLISYVGPDVNPA 1272
 Db 1207 FGGIVVLAPAKSQIFQIFFRMYLAMVLGATHGLIFLPLVLSYIGPSYKA 1258

RESULT 10
 Q59GR1 HUMAN PRELIMINARY; PRT; 1289 AA.
 AC Q59GR1;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Niemann-Pick disease, type C1 variant (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
 RA Ohara O., Nagase T., Kikuo F.R.;
 RT "None title."
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB209048; BAD92285.1; -; -.

FT NON_TER 1 1
 SQ SEQUENCE 1289 AA; 143230 MW; 663DB27C13FD801P CRC64;
 Query Match 34.8%; Score 2402; DB 2; Length 1289;
 Best Local Similarity 39.7%; Pred.No. 6.3e-161;
 Matches 525; Conservative 239; Mismatches 452; Indels 106; Gaps 26;
 QY 2 ABAGLRGMLLMA-----LLRLAQSEPYTHIOPGYCAFYDECE-----KNPELSG 47
 Db 3 ASALRGHSMRTARGALGALLLLCPAQVFSQ-----SCVMYGECCGIAYGDKRYNCEYSG 57
 QY 48 SLMTLSNVSCLSNTPARKITGDHLLILQKICRLTYTGPNTQACCSAKQLVSLSEASITK 107
 Db 58 -----PPKPLPKDGYDLVQLCELPFFFG-NVSLCCDVRQLQTLKDNLLQLPL 102
 QY 108 ALLTRCPACSDNFVNLHCHNTCSPNOSLFINVTR---VAQLGAGOLPAVVAEAPYQHS 163
 Db 103 QFLSRCPSPCFYMLLNLFCBLTCSQRSQPLNVTATEDYVDPVNTQTKTNKBLQYVVGOS 162
 QY 164 FAEQSYDSCSRVVRPAATLAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLIDIT--- 219
 Db 163 PANAMYNACRDVEAPSSNDKALGLLGGKADA-CNATNWIEMENKONGQAPFTITPVS 221
 QY 220 -PHLLEPGQAVSGIQLPNEGVARCNESQGDVATCSQDCASCPAIARPO----- 270
 Db 222 DPFV-----RGMEPMNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPMTI 273
 QY 271 -ALDSTFVLGQPGSLVLIILCSVPAVVTILL-----VGRVAPARDKSKM 316
 Db 274 LGLDAMTVIMTYMAFLVFPFGAPFVWCYRKRYFVSEYTPIDSNIAFSV-NASDKGB- 331
 QY 317 VDPKKGTSLSLDSFSTHTLQCFQGGTWTWASWPLTILVLSVPPVVALAAGLVFTLT 376
 Db 332 -----ASCCDPVSAAPEGCLRLTRWGSFCVRNPGCVIPFSLVPIACSSGLVFRVT 385
 QY 377 TDPVELMAGPNSQASEKAFHQHFGPFRFTNQVILTAPNRSYRYDSLLLLGPK-NFSGI 435
 Db 386 TNPVDLMSAPSSQARLEKEYFDQHFQFPFRTEQLIIRAPLTDKIYQYPSGADVFPFGP 445
 QY 436 LDLDLLELLELLEQLRLHLQWSPEAQRNISLQDIQYAPLNPNTSLYDCCINSLLQYEQ 495
 Db 446 LDQIQLHQVLDLQIAIEN--ITASVDNETVTLLQDICLAPLSYNT---NCTLSLVNLYQ 500
 QY 496 NNRTLLLTANTLGMQTSQVDWKDHFYCANAPITFKDGTALALSCMADYGAVPFPLA 555
 Db 501 NSHSVLDHKKGDDPF---VYADYHTHFLYCVAPASLNDTSLHDFCLGTGCGPVPV 557
 QY 556 IGGYKGYSEABALIMTFSLNYPAGDPRLAQAKLWEBAFLMEARAFORMAGMPOVTP 615
 Db 558 LGGYDDQYNNATALVITFPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTISF 614
 QY 616 TAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSRSMVMDSKATGLGGVAV 675
 Db 615 TAERSIEDELNRESDSVFTVVISYAIMELYISLALGHMKSCRLLVDSKVLGIAGILI 674
 QY 676 VLGAVMAAGPFSYLGIRSSVILQVWPPLVLSVGDNIFIVLEYQRLPRPGEPRVH 735
 Db 675 VLSVACSISGVSPYIGLPLTVIEVIFPLVAVGVDNIFILVQAYQYORDERLQGETLDQ 734
 QY 736 IGRALGRVAPSMILCSLSEAIICFFGALTPMAVRTFALTSGLAVILDFLLQMSAPVALL 795
 Db 735 LGRVLGEVAPSMFLSSFSSTVAFFLALSVMPAVHTFSLFAGLAVFIDFLLQITCFVSL 794
 QY 796 SLDSEKQASRLDVCCVKPQLPPGQ--GEGLLGFPQKAVAPFLLHWTITRGVLLFL 854
 Db 795 GLDIKQKKNRLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPILLKDMWRPIVIAIF 854
 QY 855 ALFGVSLYSMCHISVGLDQELALPKDSYLLDFLPLNRYFEVGAUVYFTTLGYNPSS 914
 Db 855 GVLSPSIAVLNKVDIGLQSLMPDDSDYMDVFKSISQYLHAGPPVYFVLEBHDYTSK 914
 QY 915 GMAIACSSAGCNFSTQIKQVATEPPEQSYLAIPASSWVDDPIDWLT--SSCCRLYISG 973

Db 915 GQMVCGGCGNNDLSVQOIFNAQAQNDYTRIGFAPSSWIDDYDWDVWKPQSSCCRV---D 971

Qy 974 PNKDFCFSTVNSLNCNCKMSIT-MGSVRPSVEQFKYLFWFLNDRPNKICPKGGAA 1032

Db 972 NITDQFCNAAVDPACVR-CRPLTEGKQRPQGGDFMFLWFLSDNPNKCGKGAA 1030

Qy 1033 STSVN--LTSQGVLASPMYHKLKNSQDYTEALRAARELANITADLRKVPSTDPAP 1090

Db 1031 SSANILLGHGTRVGATVMTYHTVLOTSDAFIDALKKARLTASNV-ETWINGS--AY 1087

Qy 1091 EYFPTITNVFEQVLTILPEGLWMLSLCLVPTEAVSCLLGLDLSRGLLNLSTVILV 1150

Db 1088 RVFVSFVVFVEQVLTIDITIFNLGVSGLAIFLVWVLGCELWSAVIMCATIAMVLV 1147

Qy 1151 DTVGFMALDISYNAVSLINLSVAGMSVEFVSHITRSFAISTKPTWLERAKBATISMG 1210

Db 1148 NMFVGMWLGISINAVSLNLSVAGMSVEFVSHITRAFTVSMKGSVERAEEALAHMS 1207

Qy 1211 AVFAGVAMTNLPGILVLGLAKAQLQIIFFRNLNLTILGLLHGLVFLPVLVYVGPVN 1270

Db 1208 SVFSGITLTKFGIGVVLVAFKSIQIFYFRMYLAMVLLGATHGLIFLPLVLSYIGPSV 1267

Qy 1271 PA 1272

Db 1268 KA 1269

RESULT 11

Q7TMD4_MOUSE

ID Q7TMD4_MOUSE PRELIMINARY; PRT; 1277 AA.

AC Q7TMD4; 25, Created

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE Niemann Pick type C1.

GN Name=Npcl;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Ruterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RL and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC052437; AH52437.1; -, mRNA.

DR EMBL; BC054539; AH54539.1; -, mRNA.

DR GO; GO:0006897; P: endocytosis; IMP.

DR InterPro; IPR004765; NP_C type.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR000731; SSD 5TM.

DR Pfam; PF02460; Patched_1.

DR TIGRFAMs; TIGR00911; ZAO60601; 1.

DR PROSITE; PS0156; SSD; 1.

SQ SEQUENCE 1277 AA; 142882 MW; 3B42230AAC8E564E CRC64;

Query Match 34.7%; Score 2400; DB 2; Length 1277;

Best Local Similarity 40.4%; Pred. No. 8.7e-161;

Matches 528; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

Qy 11 LMAILLRLAOSPEYTHIQPGYCAFYDEGKQNPGLSGLMTLSNVSCLSNTPARKITGDH 70

Db 8 LGLLLLLLCPAQVFSQ-----SCVWYGECC---IATGD---KRYNCKYSGPPKPLPKDG 55

Qy 71 LILQKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNFVNLHCHMTCS 130

Db 56 YDLVQELCPGLFP-DNVSLLCCDIQQLTKENLQPLQFLSRCPSCFYNLMTLFECLTCS 114

Qy 131 PNQSLFINVTRVAQLGAGQLPA-----VVAYEAFYQHSFAQSQSYDSCSRVRVPAATLAVG 186

Db 115 PQSQPLNVTATEDYFDPKTKQENTNKELEYFVGQSFANAMYNACRDVEAPSSNEKALG 174

Qy 187 TMCVGYSGALCNAQRNLNFQDGTGNGLAPLDI-----TFHLLPEGQAVSGIQLNEGVA 241

Db 175 LLCGRDARA-CNATNWIEMFNKNGQAPFTIIPVFSDSLIL-----GMEPMRNATK 225

Qy 242 RCNCSQGDVATCSQDCQCAASC-----PATARPOALDSTF-----YLGQMPG 283

Db 226 GCHESVDEVTGPCSCQDCSIVCGPKPQPPPPMPKRWINGLDAMYIMVTVYVFLFVFFG 285

Qy 284 SILVILILCSFVAVVITLLVGRFVAPARDKSMVDPKGTSLSKLSKLSSTHTLLGQFTQG 343

Db 286 ALLAVWCHRRYFVSEYTPIDSNIAFVNS-----DKGEASCCDPLGAFFDDCLRMFTK 341

Qy 344 WGTWASWPLTILVLSVTPVVALAAGLYFTLTTPVELWSAPNSQARSEKAFHDHFGP 403

Db 342 WGAFCVRNPTCIIFSLAFITVCSGLVFGVQVTPNVELWSAPHQARLEKEYPKHFGP 401

Qy 404 RPTNQVILTAPNRSYRYSVLSLLGPK-NFSGILDLLELLELLELLELLELLELLELLEL 462

Db 402 FRTEQLIIQANTSVHLYEYPYAGADVPFGPPLNKEILLHQVLDLQ-----IAIESITAS 456

Qy 463 RN---ISLQDICYAPLNPTNTSLYDCCINSLLQYFQNNRTLLLTANTQTLMGQTSQV--- 516

Db 457 YNNETVTLQDICVAPLSPYK---NCTIMSVLNYFQNSHAVL-----DSQVGD 502

Qy 517 -----DWKDHFLYCANAPLTFKDGITALALSCWADYGAPEFLAIGYKGDYSAEALI 571

Db 503 FYIYADYHFLYCYVRAPASLNDTSLHGPCUGTGTGGVPVFWLGGYDDQNNYNTALV 562

Qy 572 MTFSLNNYPAGDPRLAQAALWEAELEEMRAFQRMAGMFQVTTAERSLEDEINTTAE 631

Db 563 IIFPVNNYNDTERLQRAWAKEFEISFVKYKN---PNLTISFTAESRTEDBELNES 619

Qy 632 DLPIFATSYIVFLYISLALGYSWSRVWVDSKATLGLGVAVNLGAVMAAGFFSYLG 691

Db 620 DVFTVIISYVWVFLYISLALGHIQSCSRLVDSKISLGIAGLIVLSSVACSLGIFSYMG 679

Qy 692 IRSSLVILQVWPLVLSVGADNIFIVLEYQRLPRRPGEPREVIHIGRALGRVAPSMLLCS 751

Db 680 MPLTLIVIEVIFPLVLAAGVDNIFILVQYQDERLQETLDQQLGRILGEVAPTMFLSS 739

Qy 752 LSEAI CFFLGA LTPM PAVRTFALTSG LAVILFLQMSAFVALLSDLSQKQASRLDVCC 811

Db 740 FSETSAFFFGALSSMPAVHTFSLFAGMAVLIDPLQLQITCFVSLGLDIDIKRQKNHLDILC 799

QY	1112	GLFMLSCLVPTFAVSCLLGLDLSRGLNLLSIWMILVDTVGMALWDISYNAVSLNL	1171
Db	1096	TIFNLSVSLGAIFLVTLLGCDLWSAIVMCIITAMILVNMFGVMWMLGSLNAVSLNL	1155
QY	1172	VSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAK	1231
Db	1156	VMSCGISVEFCSHITRAFTVSMKSGRAQRAEALAHMGSSVFGSITLTKFGGIVVLAPAK	1215
QY	1232	AQLIQIFPRNLITLLGLLHGLVFLPVILSVYGVDPVNPALAL---EQKRAE 1282	
Db	1216	SQLFQIFYFRMYLAWVLGATHGLIFLPLVLLSYIGPSINKAKSLATQEQYKGTGTE 1269	
RESULT 13			
Q8M149_FELCA	ID Q8M149_FELCA PRELIMINARY; PRT; 1276 AA.		
AC	Q8M149;		
DT	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	Mutant Niemann-Pick C1.		
OS	Felis silvestris catus (Cat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;		
OC	Felis.		
OX	NCBI_TaxID=9685;		
RI	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22695580; PubMed=12809639; DOI=10.1016/S1096-7192(03)00074-X;		
RA	Somers K.L., Royals M.A., Carstea E.D., Rafi M.A., Wenger D.A.,		
RA	Thrall M.A.;		
RT	"Mutation analysis of feline Niemann-Pick C1 disease.";		
RL	Mol. Genet. Metab. 79:99-103(2003).		
DR	EMBL; AF503633; AAM27450.1; -; mRNA.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0005764; C:lysosome; IEA.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008158; F:hedgehog receptor activity; IEA.		
DR	GO; GO:0030301; P:cholesterol transport; IEA.		
DR	InterPro; IPR004765; NP_C-type.		
DR	InterPro; IPR003392; Patched.		
DR	InterPro; IPR000731; SSD 5TM.		
DR	Pfam; PF02460; Patched; 1.		
DR	TIGRFAMs; TIGR00917; 2A060601; 1.		
DR	PROSITE; PS50156; SSD; 1.		
SQ	SEQUENCE 1276 AA; 141694 MW; CB3562214195998F CRC64;		
Query Match 34.7%; Score 2395; DB 2; Length 1276;			
Best Local Similarity 40.6%; Pred. No. 2e-160;			
Matches 534; Conservative 232; Mismatches 462; Indels 86; Gaps 29;			
QY	7	RGWLWALLRLAQSBEYTHIQGYCAFYDECGKNPGLSGMLTSLNVSCLSNTPARKI 66	
Db	4	RGPALGLLLLC---PAQVLAQS--CIWYGECC---IASGD---KRYNCKYSGPKPL 51	
QY	67	TGDLHLLIKICPLRYTGNTQACCSAKQLVLEASLSITKALLTRCPACSDNFVNLHCH 126	
Db	52	PKQGYDLVQLCPGFFF--DNVSLCCDVQQLQTLKONLQLPLQLSRCPSCFYNLVLFCE 110	
QY	127	NTCSPNQSLFINVTR----VAQLGAGQLPAVVAEAFYQHSFABQSDYSCSRVVPAAAT 182	
Db	111	LTCSPRQSQFLNVATATEDYDVPVNTQKTNVKELQYYIGESFANAMYNACRDVEAPSSND 170	
QY	183	LAVGTMCGVYGSALCNARQMLNFOGDTGNGLAPLDIT--PHLLEPGQAVSGIQPLNEGV 240	
Db	171	KALGLLCKGDAEA--CNATNMI EYMFNSKDNQGAFFTITPISDLPT-----HGMPEMNAT 224	
QY	241	ARCNESQGDVATCSCQDCAASC-----PAIARPAQLDSTFYLGMPGSLVLIIL 291	
Db	225	KGCDSESDVEVTGPCSQDCSIVCGPKPQPPPPVFWRLGLDAMVIMWITYMAFLVFF 284	
QY	292	CSVFA-----VVTILLVGRVA---PARDKSKMDPKGTSLSDKLSFSTHTLLGQ 339	

Db	285	GAPFALWCYKRYFVSEYTPIDSNIAFVSNDRGE-----ASCCDALGAFAECCLRR 337	
QY	340	FFCGWTWASWPLTTLVLVSVVVALAAGLVFTELTTPDVELWSAPNSQAREKAFHQ 399	
Db	338	LFSQWGSFCVRNPGPIFFSLAFIACSSGLVFRVTTNPVDLWSAPSSQARLEKEYFDT 397	
QY	400	HFGPFRRTQVILTAPNRSSRYDLSLLGPK-NFSGILDLLLELLELERLHQLQVMS 458	
Db	398	HFGPFRTEQLIIQAHTSAHTYQYPYSGSDVFPGLDLAILHQVLDLQTAIEN--ITA 455	
QY	459	PEAQRNISODICYAPLNPDTSLYDCCINSLLQYQNNRTLLLTANQTLMGQTSOV-- 516	
Db	456	SYNNETVTLQIDICVAPLSPYK---NCTILSVLNFQNSHML-----DHEIGDFFVYA 507	
QY	517	DKWDHFLYCANAPLTPKDGFTALALSCWADYGAVPFPFLAIGYKGYKDYSAEALIMTFSL 576	
Db	508	DYTHLLYCVRAPASLNDTSLLDHPCLGTFGGVFPVPLVLGGYDDQNNYNATALVITFPV 567	
QY	577	NNYPAGDPRLAQAQKLEEAFLSEEMRAFORMAGHQVTTAERSLEDEINRTTAEDLPF 636	
Db	568	NNYNDTERLQKAHVWEKEFINFVNYKN---ENLITISFTTTERSIEDELNRESNGDFTV 624	
QY	637	ATSVIVFLYISLALGSYSWSRVMDSKATLGLGVAVVLGAVMAAGPFYSIGIRSSL 696	
Db	625	IISYAIMFLYISIALGHKISCRLLVDSKISLGIAGLILVLSVACSLGIFSIVGIPLYL 684	
QY	697	VILQVPPFLVSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAI 756	
Db	685	IVIEVIFPLVAVGVNDFILVQTYQDERLHGETLQQLGRVLGEVAPSMFLSSPSEAV 744	
QY	757	CFFLGALTPMPAVRTFALTGLAVILFLOMSAFVALLSDSKRQASRLDVCCKVQP 816	
Db	745	AFFLGALSMPAVHTFSLFAGMAVLIDFLQITCFVSLGLDIDIKRQEKRLDVLCCVRGS 804	
QY	817	ELPPPGQ-GEGLLGGFFQKAYAPELLHWITRGVLLFLALFGVLSVSMCHISVGLDOEL 875	
Db	805	EDGTSVQASECLFLFKHSYSPLLDKWMPVIAIFVGLVSFAVLNKVEIGLDQSL 864	
QY	876	ALPKDSYLLDYFLNRYFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCANNFPTQIKI 935	
Db	865	SMFDDSYVMDYFKSL-KYLHAGPPVYFVLEBHDYTSLKGNQVMVCGMGCNNDLSVQQIF 923	
QY	936	YATEFPQSYLAIPASSWDDFDWLTP-SSCCRLYISGPNKDKFCPSVNSLNLCKNM 994	
Db	924	NAAQDSYTRIGFAPSSWIDDYFDWVKPQSSSCRYVNS---TDRPCNASVVDPAICR-CR 979	
QY	995	SITM-GSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN-LTSDGOVLASRFMAY 1052	
Db	980	PLTQEGKORPOGGDFMFLPMFLSDNPNPKCGKGHAAVSSAVNLGNDTGVGATYFTY 1039	
QY	1053	HKPLKNSQDYTEALRAARELAANITADLRKVPGTDDPA-FEVPYTYITVTFEQLTILPE 1111	
Db	1040	HTVLQTSADFTDAMERKANLIASNIT---KTMLEGSNRYRVPYFVVFYEQYLTIID 1095	
QY	1112	GLFMLSCLVPTFAVSCLLGLDLSRGLNLLSLIWMILVDTVGMALWDISYNAVSLNL 1171	
Db	1096	TIFNLSVSLGAIFLVTLLGCDLWSAIVMCIITAMILVNMFGVMWMLGSLNAVSLNL 1155	
QY	1172	VSANVMSVEFVSHITRSFAITKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAK 1231	
Db	1156	VMSCGISVEFCSHITRAFTVSMKSGRAQRAEALAHMGSSVFSGITLTKFGGIVVLAPAK 1215	
QY	1232	AQLIQIFPRNLMLITLLGLHLGVFLFVILSVYGVDPVNPALAL---EQKRAE 1282	
Db	1216	SQLFQIFYFRMYLAWVLGATHGLIFLPLVLLSYIGPSINKAKSLATQEQYKGTGTE 1269	
RESULT 14			
ID	NPCL_MOUSE	STANDARD; PRT; 1278 AA.	
AC	Q35604	Q35604; Q35605;	
DT	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		

10-MAY-2005 (Rel. 47, Last annotation update)
 Niemann-Pick C1 protein precursor.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 NUCLEOTIDE SEQUENCE.
 MEDLINE=97362324; PubMed=9211850; DOI=10.1126/science.277.5323.232;
 Loftus S.K., Morris J.A., Carstee E.D., Gu J.Z., Cummings C.,
 Brown A., Ellison J., Ohno K., Rosenfeld M.A., Tagle D.A.,
 Pentchev P.G., Pavan W.J.;
 "Murine model of Niemann-Pick C disease: mutation in a cholesterol
 homeostasis gene.";
 Science 277:232-235 (1997).
 [2]
 SUBCELLULAR LOCATION.
 MEDLINE=99145606; PubMed=9990080; DOI=10.1073/pnas.96.4.1657;
 Patel S.C., Suresh S., Kumar U., Hu C.Y., Cooney A.,
 Blanchette-Mackie E.J., Neufeld E.B., Patel R.C., Brady R.O.,
 Patel Y.C., Pentchev P.G., Ong W.-Y.;
 "Localization of Niemann-Pick C1 protein in astrocytes: implications
 for neuronal degeneration in Niemann-Pick type C disease.";
 Proc. Natl. Acad. Sci. U.S.A. 96:1657-1662 (1999).
 -!- FUNCTION: Involved in the intracellular trafficking of
 cholesterol. May play a role in vesicular trafficking in glia, a
 process that may be crucial for maintaining the structural and
 functional integrity of nerve terminals.
 -!- SUBCELLULAR LOCATION: Integral membrane protein. Found in late
 endosomes and lysosomes.
 -!- TISSUE SPECIFICITY: Expressed predominantly in perisynaptic
 astrocytic glial processes. Also expressed in heart, spleen, lung,
 liver, skeletal muscle, kidney, testis.
 -!- INDUCTION: Activated by the drugs progesterone and U-18666A which
 block cholesterol transport out of lysosomes and by the
 lysosomotropic agent NH4Cl.
 -!- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain
 containing a di-leucine motif necessary for lysosomal targeting
 are critical for mobilization of cholesterol from lysosomes.
 -!- SIMILARITY: Belongs to the patched family.
 -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

 EMBL; AF003348; AAB63372.1; -; mRNA.
 EMBL; AF003349; AAB63373.1; -; Genomic_DNA.
 PIR; T30188; T30188.
 Ensembl; ENSMUSG0000024413; Mus musculus.
 MGI; MGI:109712; Npcl.
 GO; GO:0006897; P: endocytosis; IMP.
 InterPro; IPR004765; NP_C type.
 InterPro; IPR003392; Patched.
 InterPro; IPR000731; SSD_5TM.
 Pfam; PF02460; Patched; 1.
 TIGRFAMs; TIGR00917; 2A060601; 1.
 PROSITE; PS0156; SSD; 1.
 Glycoprotein; Lysosome; Signal; Transmembrane.
 FT SIGNAL 1 23 Potential.
 FT CHAIN 24 1278 Niemann-Pick C1 protein.
 FT TRANSMEM 271 291 Potential.
 FT TRANSMEM 352 372 Potential.
 FT TRANSMEM 623 643 Potential.
 FT TRANSMEM 655 675 Potential.
 FT TRANSMEM 685 705 Potential.
 FT TRANSMEM 761 781 Potential.
 FT TRANSMEM 834 854 Potential.
 FT TRANSMEM 1099 1119 Potential.

FT TRANSMEM 1125 1145 Potential.
 FT TRANSMEM 1196 1216 Potential.
 FT TRANSMEM 1228 1248 Potential.
 FT DOMAIN 621 786 SSD.
 FT MOTIF 1275 1278 Di-leucine motif.
 FT COMBIAS 250 260 Poly-Pro.
 FT CARBOHYD 71 71 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 123 123 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 138 138 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 186 186 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 223 223 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 229 229 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 415 415 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 460 460 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 479 479 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 525 525 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 474 474 S -> F (in Ref. 1; AAB63373).
 FT CONFLICT 479 479 N -> D (in Ref. 1; AAB63373).
 SQ SEQUENCE 1278 AA; 142889 MW; 43C71CE47D283674 CRC64;
 Query Match 34.5%; Score 2385; DB 1; Length 1278;
 Best Local Similarity 40.2%; Pred. No. 1e-159;
 Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;
 QY 14 LLLRLAQSEPTTIHQGYCAFYDEGKNPELSGLMTLSNVSLSTNPARKITGDHLIL 73
 DB 12 LLLLLCPAQVFSG-----SCWYEGCG---IATGD---KRYNCKYSGPPKPLPKDGYDL 59
 QY 74 LQKICPRLTYGTNTQACCSAKOLVSEASLSITKALLTRCPACSDNFVNLHCINTCSPNQ 133
 DB 60 VQELCPGLFF-DNVSLLCCDIQQLTKSLQLPLQFLSRCPSCFYNLMTLFCBLTCSPHQ 118
 QY 134 SLFINVTRVAQLGAGQLPA---WVAYEAFYQHSFAEYSQSDSCSRVVPAAATLAVGTWC 189
 DB 119 SQFLNVTATEDYFDPTKNTVKELEYVQCSFANAMYNACRDVEASSNEKALGLLIC 178
 QY 190 GYVGSALCNQARWLNFQGDGTGNGLAPLDI-----TFHLEPQAGVSGIQPLNMGVARCN 244
 DB 179 GRDARA-CNATWIEWFMKONGQAPFTIIPVPSDLSIL-----GMEPMENATKCN 229
 QY 245 ESQGDVATCSQDCNASC-----PAIARPOALDSTP-----YLQMPGSLV 286
 DB 230 ESDVEVTGSCQDCSIVCGPKPQPPPPMPWRIMGLDAMYIMWYVYAFVFPFGALL 289
 QY 287 LLIILCSFVAVVTILLVGRVAPADKSKMVDPKGTSLSDKLSPTHTLLIGOFFQSGWT 346
 DB 290 AWCCHRRYFVSEYTPIDSNIAFSVNSS-----DKGEASCCDPLGAADFCLRMFTKGA 345
 QY 347 WVASWPLTILVLSVIPWVALAAGLVFTLTTPDELWSAPNSQARSEKAFHQHGFPPFR 406
 DB 346 FCRNPTCIIPFSLAFITVCSGLVQVVTNPVELWSAPHQARLEKEYFDKHFPPFR 405
 QY 407 TNQVILTAPNRSSRYDSLLGPK-NFSGILDLDDLELLELLELLELLELLELLELLELLEL 464
 DB 406 TEQLIQAPNTSVHTYBYPAGADVPFGPLNKEILHQLVNLQ-----IAISITASYNN 460
 QY 465 --ISLQDICYAPLNPDNTSLYDCCNSLSLQYFQNNRTLLLTANQLTLMQTSQV----- 516
 DB 461 ETVTLQDICVAPLSYFNK-----NCTIMSVLNLYFQNSHAVL-----DSQVGDFFY 506
 QY 517 --DWKDFLYCANAPLTFKDGDTALALSCMADYGAIVFPFPLAIGYGYKDYSAEALIMTF 574
 DB 507 YADYHTFLYCVRAPASLNDTSLHGPCLGTGSGVFPVFWLVGLGYDDQNNYNATALVITP 566
 QY 575 SLNNYPAGDPLRAQAKJWEBAFLBEMRAPQRMAGMQVPTFTABRSLDEINRTTADLP 634
 DB 567 PVNNTYNDTERLQRAWAKEKEISFVKYKN---ENLTISFTABRSIEDELRESNSDVP 623
 QY 635 IPATSVIVFLYISLALQSYSSWSVMVDKATLGLGGVAVVLGAVMAWAGFYSYLGIRS 694
 DB 624 TVIISYVMFLYISLALGHQISCRLLDLSKISLIAGILIVLSSVACSLGIFSYNGMPL 683
 QY 695 SILVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHTIGRALGRVAPSMLLCSLSE 754


```
Db      |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
994  DFMRFPLMFLSDNPNPKCGKGGHAAYSAAVNILDNGTRVGATYFMTYHTVLOTSAADFIDA 1053
QY      :  |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1066  LRAARELAANITADLRKVPCTDPAPEVFPVTITNVFEQVLTILPEGLFMLSCLVPTFA 1125
Db      :  |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1054  MEKARLIASNITFTMQQCGD---HRVPYSVPYVEQVLTIMIDDTIFNLSVSLGAIFL 1110
QY      |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1126  VSCLLGLDLRGLNLLSIVMILVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHI 1185
Db      |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1111  VAVVLGCBLSAVIMCATIAMILVNMFGVMWLGISLNAVSLVNLVMSCGISVBFCSHI 1170
QY      |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1186  TRSFAISTKPTWLERAKEATISNGSAVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLL 1245
Db      |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1171  TRFTVSTTKGSRVERAEALSHMGSSVFSGITLTKFGGIIVLAPAKSQIFQIFYFRMYLA 1230
QY      |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1246  ITLLGLLHGLVFLPVILSVYVGDVNPALAL 1275
Db      :  |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:| |.:|
1231  MWLLGATHGLIFLPLVLLSYIGPSINKAKSL 1260
```

Search completed: April 7, 2006, 19:21:05
Job time : 274 secs

THIS PAGE IS BLANK

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:21:19 ; Search time 52 Seconds
(without alignments)
2117.769 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGRGWLLWALLRLAQ.....GSIKAGALSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/taa/5 COMB.pep.*
2: /cgn2_6/prodata/1/taa/6 COMB.pep.*
3: /cgn2_6/prodata/1/taa/H COMB.pep.*
4: /cgn2_6/prodata/1/taa/PCUTS COMB.pep.*
5: /cgn2_6/prodata/1/taa/RE COMB.pep.*
6: /cgn2_6/prodata/1/taa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2403	34.8	1318	2	US-09-949-016-10152
2	2402.5	34.8	1278	2	US-09-462-136-2
3	2385	34.5	1319	2	US-09-462-136-4
4	1329	19.2	1170	2	US-09-462-136-6
5	1078	15.6	1296	2	US-09-462-136-9
6	651.5	9.4	1447	1	US-08-540-406-19
7	651.5	9.4	1447	2	US-08-656-055-19
8	651.5	9.4	1447	2	US-08-954-668-19
9	651.5	9.4	1447	2	US-09-268-140-5
10	651.5	9.4	1447	2	US-08-918-658-19
11	651.5	9.4	1447	2	US-09-724-631-19
12	651.5	9.4	1447	2	US-08-954-701A-19
13	651.5	9.4	1447	2	US-09-807-007-6
14	651.5	9.4	1447	2	US-09-754-032-19
15	651.5	9.4	1447	2	US-08-916-140-19
16	651.5	9.4	1447	4	PCT-US95-13233-19
17	646.5	9.4	1434	1	US-08-540-406-10
18	646.5	9.4	1434	2	US-08-656-055-10
19	646.5	9.4	1434	2	US-08-954-668-10
20	646.5	9.4	1434	2	US-08-918-658-10
21	646.5	9.4	1434	2	US-09-724-631-10
22	646.5	9.4	1434	2	US-08-954-701A-10
23	646.5	9.4	1434	2	US-09-754-032-10
24	646.5	9.4	1434	2	US-08-916-140-10
25	646.5	9.4	1434	4	PCT-US95-13233-10
26	639.5	9.3	642	2	US-09-248-796A-16161
27	628	9.1	1296	2	US-08-857-636-60

28	584	8.5	1203	2	US-09-293-505-2	Sequence 2, Appli
29	584	8.5	1203	2	US-09-060-939A-2	Sequence 2, Appli
30	583.5	8.4	1182	2	US-09-293-505-7	Sequence 7, Appli
31	583.5	8.4	1182	2	US-09-060-939A-7	Sequence 7, Appli
32	581.5	8.4	531	2	US-09-248-796A-16160	Sequence 16160, A
33	580	8.4	1203	2	US-09-207-857-2	Sequence 2, Appli
34	580	8.4	1203	2	US-09-909-280A-2	Sequence 2, Appli
35	579	8.4	1146	2	US-09-807-007-1	Sequence 1, Appli
36	526.5	7.6	1311	1	US-08-540-406-4	Sequence 4, Appli
37	526.5	7.6	1311	2	US-08-656-055-4	Sequence 4, Appli
38	526.5	7.6	1311	2	US-08-954-668-4	Sequence 4, Appli
39	526.5	7.6	1311	2	US-08-918-658-4	Sequence 4, Appli
40	526.5	7.6	1311	2	US-09-724-631-4	Sequence 4, Appli
41	526.5	7.6	1311	2	US-08-954-701A-4	Sequence 4, Appli
42	526.5	7.6	1311	2	US-09-754-032-4	Sequence 4, Appli
43	526.5	7.6	1311	2	US-08-916-140-4	Sequence 4, Appli
44	526.5	7.6	1311	4	PCT-US95-13233-4	Sequence 4, Appli
45	474	6.9	1286	2	US-09-268-140-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-10152
; Sequence 10152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10152
; LENGTH: 1318
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10152

Query Match 34.8%; Score 2403; DB 2; Length 1318;
Best Local Similarity 39.7%; Pred. No. 3.9e-218;
Matches 525; Conservative 238; Mismatches 451; Indels 110; Gaps 26;

QY	2	ABAGRGWLLWA-----LLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSG 47
DB	32	ASAAALRGHSMARGALGLLLLLCPAQVFSQ-----SCVWYGECCYAYGDKRYNCEYSG 86
QY	48	SLMTLSNVSCLSNTPARKITGDHLLLOKICPLRYTGNPTQACCSAKQLVSLASISITK 107
DB	87	-----PPKFLPKDGYDLVQLCPLGPFEG-NVSLCCDVRQLQTLKDLQLPL 131
QY	108	ALLTRCPACSDNFVNLHCHNTCSPNQSLFINVTR-----VAQIAGAGQLPAVVAEAFYQHS 163
DB	132	QFLSRCPSCFYNNLNFCELTCSPPSQFLNVATEDYDPTVNTQTKTNKELQYVVGOS 191
QY	164	FAQSYDSCSRVRPAATLAVGTWCGVYGSALCNAQKWLNFQDGTGNGLAPLDIT-----219
DB	192	FANAMYNACRDVEAPSSNDKALGLLCGKDADA-CNATNMIETMFKNDQCAPFTITPVFS 250
QY	220	---FHLLEPGQAVGSGIOPLENGVARNCHESQGDVATCSQDCAASCAPARPO-----270
DB	251	DFPVH-----GMEPFMNATKGCDESVDVETAPCSCQDCSIVCGPKPQPPPPPPAW 300
QY	271	---ALDSTFYLGQMPGSLVLIILILCSVFAVVTILL-----VGPRVAPARDKS 314

Qy	1162	SYNAVSLINLVAVGMSVEFVSHITRGSFAISTKPTWLERAKEATISMGSAVFAGVAMTNL	1221
Db	1189	SLNAVSLVNLVMSGIVSEFCSHITRAFTSTKGRSVRAEEALAHMGSSVFSGITLTKF	1248
Qy	1222	PGILVLGLAKAQLIQIFFRNLNLITLGLLHGLVFLPVLSYVGPDPNPA	1272
Db	1249	GGIVLAPAKSQIPEIFRMYLAMVLGATHGLIFLPLVLLSYIGPSVANKA	1299
RESULT 4			
US-09-462-136-6			
; Sequence 6, Application US/09462136			
; Patent No. 6426198			
; GENERAL INFORMATION:			
; APPLICANT: Carstee, et al.			
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease			
; FILE REFERENCE: 4239-53894			
; CURRENT APPLICATION NUMBER: US/09/462,136			
; CURRENT FILING DATE: 2000-06-01			
; PRIOR APPLICATION NUMBER: PCT/US98/13862			
; PRIOR FILING DATE: 1998-07-02			
; PRIOR APPLICATION NUMBER: US 60/051,682			
; PRIOR FILING DATE: 1997-07-03			
; NUMBER OF SEQ ID NOS: 13			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 6			
; LENGTH: 1170			
; TYPE: PR1			
; ORGANISM: Saccharomyces cerevisiae			
US-09-462-136-6			
Query Match			
Best Local Similarity 28.8%; Pred. No. 3.4e-116;			
Matches 373; Conservative 230; Mismatches 512; Indels 182; Gaps 46;			
Qy	10	LLWAL-----LRLAQSBYTTIHQPGYCAFYDECQNP-----ELSGSLMTLSNVCSLSN	60
Db	3	VLMIIALVGLMRLVQG-----TATCAMYGNCGKSKSVFNGELPCVPFRSFEPPVLSD	54
Qy	61	TPARKITGDHLLIQLKICPRLYTGPNTOACCSAKOLVSLASLITKALLTRCPACSDNF	120
Db	55	ETSK-----LLVEGCGEWEKVR-YACCTKDQVVALRDNLQKAPLISPCACLKNF	105
Qy	121	VNLHCHNTSPNQSFLNTRVAQLGAGOLPVAAYEAFYQHSFASQSDYSCSRVRPAA	180
Db	106	NNLFCHFTCAADQGRFVNITKV-EKSKEDKDI VAEILDVFNSSWASEFYDCKNIKFSAT	164
Qy	181	ATLAVGTMCVGYGALCNAORWLNFGQDTGN--GLAPLDITFHLLEPGQAVSGIQLNE	238
Db	165	N----GYAMDLIGGAKNYSQFLKFLGDAKPMLGSGSPQINIKYKIDLANEE--KEWQEFND	218
Qy	239	GVARNESQDDVATCSQDCQCAASCPAIAAPQALDSTFYLGOMFG---SLVLIILCSVF	295
Db	219	EVVACDAQ---YKACSCQSCPHL-KP-LKDGCKVKGVLPCFSLSLVLIPTICALF	272
Qy	296	AVVTILVGRVAPARDKSMVDPKGTSLSDK-----LSFSTHT-----LLQPFQCGW	345
Db	273	AFMYYLCKRKNGAMIVDDDIVPESG-SLDESETNVFESFNNETNPFNGKLANLFTKV	331
Qy	346	TWASWELTILVSVIPVALAAGLV-FTLETTPVELWSAPNSQARSEKAFHQHGPFF	404
Db	332	QFSVENPKILITTVFSIFVFSFIQVATLETDPINLWSSKSEKSEKEYEDDNFGPF	391
Qy	405	FRTHQVILTAPNRSSRYDLSL--LLGPKNFSGIILDLDLLELLEBLRHLQWSPSAQ	462
Db	392	YRTQIFVNVETGPVLSVETLHWFDVENF-----ITEEL-----QSS	429
Qy	463	RNISLQDTCYAPLAPNDNTSLVDCINSLLQVFQNNRLLLLTANQTLMGQTSQVDWKDH	522
Db	430	ENIGYQDLCPRP-TEDST-----CVIESFTQYFOG-----ALPNKDSWKREL	470
Qy	523	LYCANAPLTFKDGITALCSMDYGPVFPPELGIYKKGKDYSEAEALIMTFSLNNYPAG	582
Db	471	QECGKFP-----VNCLPTFQOPLKTNLL---FSDDILNAHAFVVVLTLLTNH---	514
Qy	583	DPLAQAKWERAFLEEMRAFORRMAGMFQVFTTAERSLEDEINRTABDLPFATSYIV	642
Db	515	---TQSANWEER-LESEYLLDLKVPBGL-RISFNTESISLEKLNNN--NDISTVAISYLM	567
Qy	643	IFLYISALGSSYSWSRWVDSKATIGLGGVAVVLGAVMAAMGFPSYGLIRSSILVILQVV	702
Db	568	MFLYATWALRRKDGKTRLL-----LGISGLLIVLASIVCAAGFLTLFGLKSTLIIEVI	621
Qy	703	PFLVLSVGDADNIFIFVLEYOR-LPRRPGPREVHIIGRALGRVAPSMMLCSLSAICFFLG	761
Db	622	PFLILAIIDINIFLITHEYDRNCEQKPEYSIDQKIISAIGRMSPSILMSLLCOTGCFLIA	681
Qy	762	ALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKSQEASRLDVCCKVPQELPPP	821
Db	682	AFVTMPAVHNFAYISTVSIFNGVLQLTAYVLSLSLYEKRSNYKQIT-----	728
Qy	822	GQGBGLLGFQKAYAPFLHMTIRGVWLLFLALFGVSLYSMCHISVGLDQELALPKDS	881
Db	729	-GNEETKESFLKTFYFKMLTQ---KRLIIIFSNWFFTSILVFLPEIQFGLDQTLAVPQDS	784
Qy	882	YLLDYFLNRYPEVGVAPYVVTTLGVNFSSEAGMNAICSS-AGCINNFSFTQKIYATEF	940
Db	785	YLVDFYKDVYSFLNVGPPVYMWVK-NLDLTRQNOQKICGKFTTCERDLSLANVLE---QE	840
Qy	941	PEQSYLAIPASSWYDDFIDMLTPSS--CCRLYISGPNKDKFCPSTVNSLNCNKCMSTIM	998
Db	841	RHRSTITEPLANWLDYFMFLNPQNDCCRL---KKGTDVCPSPSPFSRRRC-ETCFQ--Q	894
Qy	999	GSVR-----PSVEQFHKYPFWFLNDRPNIKCPKGGAAAYSTSVNLSDGOVLASRFMAY	1052
Db	895	GSWYNWMSGFPEGKDFMEYLSIWIN-APSDPCPLGGRAPYSTAL-VYNETSVSASFRTA	952
Qy	1053	HKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITVNFYQYLITLPEG	1112
Db	953	HHPLRSQKDFQAV-----SDGVRISSPFPELDMFAYSFPFYFFVQYQTLGPLT	1001
Qy	1113	LFMLSLCLVPTFAYSCLLGLDLRSLNLASIVMILVDTVGFMALMDISYNVSLNLIV	1172
Db	1002	LKLGSAIILIFFISSVFL-QNIRSSFLALVVTMIIIVDIGALMALLGISLNAVSLNLI	1060
Qy	1173	SAVGMSVEFVSHITRGSFAI---STKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLG	1229
Db	1061	ICVGLGVFCVHIVRSFTVPVSETKCDANSRVLSLNTIGESVIKGITLTITKFIGVCVLAF	1120
Qy	1230	AKAQILQIOPFRNLNLITLGLLHGLVFLPVLSYVG	1266
Db	1121	AQSKIFDVYFRMFWFTLIIVAALHALLFLPALLSLFG	1157
RESULT 5			
US-09-462-136-9			
; Sequence 9, Application US/09462136			
; Patent No. 6426198			
; GENERAL INFORMATION:			
; APPLICANT: Carstee, et al.			
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease			
; FILE REFERENCE: 4239-53894			
; CURRENT APPLICATION NUMBER: US/09/462,136			
; CURRENT FILING DATE: 2000-06-01			
; PRIOR APPLICATION NUMBER: PCT/US98/13862			
; PRIOR FILING DATE: 1998-07-02			
; PRIOR APPLICATION NUMBER: US 60/051,682			
; PRIOR FILING DATE: 1997-07-03			
; NUMBER OF SEQ ID NOS: 13			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 9			
; LENGTH: 1296			
; TYPE: PR1			
; ORGANISM: Caenorhabditis elegans			
US-09-462-136-9			

Query Match	15.6%	Score 1078;	DB 2;	Length 1296;
Best Local Similarity	24.0%	Pred. No. 2.8e-92;		
Matches 321;	Conservative 237;	Mismatches 492;	Indels 288;	Gaps 41;
76	KICPLRYTGNTOACCSAKQVLSLEASISITKALLTRCPACSDNPNVNLCHTCSNOSL	135		
63	EFCPHLLTGDN-KLCCCTPSQAGLTKQIAQHILGRPCFDNPAKLWCSEFTCSNQOD	121		
136	FINVTRVACL--GAGQLPAVVAEAF-----YQHS--FAEQSYDSCSRVRVPAATAATLAVG	186		
122	FVISEMKEPIEGEGFTPEYQPAEAYVNTVEYRLSTDFAEGMFSCKDVTGGQPALRV-	180		
187	TWCGVYGSAICQAQWNLNFOGDTGNGL-APDITPHLEPCQAVGSGIQP-LNKGVARCN	244		
181	-MC---TSTPCTLTNMLEFIFGTQNDLNIPIHTFKLLYDPIKTPPSRSTYMNMYNFTGCD	236		
245	ESQGDVATSCQDCAASCPAIRQALDSTFYLGOM-----PGSLVLIILICSVFAVVT	299		
237	KSARVGWACSTSEC--NKEEVANLIDLDGKTSQTCNVHGIACLNIFVWLAFIGSLAV	294		
300	ILLVGF-----RVAPARDKSKMVPDKGTSLSDKLSFSTHTLLGQFF	341		
295	LLCVGFVTSYDEYDYNLRQTSQGESPKRNRIK-----RTGAWI-----HNFMENNA	342		
342	QGWGTWASWPLTILVLSVIPVALAAGLVTELTDPVELWASAPNQARSEKAFHQHP	401		
343	RDIGMAGRNKSHFFICGAVLIFCLPGMITYHKESTVNVDMWMSPRGRQBEVFNANF	402		
402	GPFFRTNQVILTAPNRSSRYRDSLLGPKNPSGILDLDLLELLEQLERLHLQVWSPA	461		
403	GRPQYQIMLL--SHRDFQSSGKLYG-----VPHKDIPEELFDILNAKNISTQSDG	455		
462	QRNLSQDICIAPLNPDNTSYDCINSLLQYFQNNRTLLLTANQT-----L	509		
456	-RTITLDVCYRPMGPG---YDCLINSPTNYFQCNKEHLDMKSKBETVSEDDDAFDYF	510		
510	MGQTSQVMDKHDELYCANAPLTFKDGATALASCMADYCAPVFPFLAIGYKGDYSAEA	569		
511	SSEATTDEWNNHMAACIDQPMQK--TKSGUSCMGTGGPSAPNN-VFGKNSNTHQANS	567		
570	LIMTFLSNYPAGDPRLLAQAKLBEAFLEENRAFQRRMAGMFQVTTAERSLEDEINRTT	629		
568	IMWTLVTQ--RTEPEIQKASLMEKEFLKFCKEYREKSPKVI-FSFWAERSITDEIENDA	624		
630	AEDLPIFATSVIVIFLYLSALGSY-----SSWRVWVDSKATGLGGAVALGANVAM	684		
625	KDEIVTVVIALAFLIGYVTFSLGRYFCVENQWS-ILVHSR-----	664		
685	GFFSVLGRISRLVQLVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHIGRALGRVA	744		
665	GF-----	666		
745	PSMLLCSLSEACFFLGALTPMPAVRTPALTSGLAVIDLFLLOMSAFVALLSDSKRQEA	804		
667	-----TDLPAIRTFCLVAGLAVLIDVVLHCTIFLALFWDTQRELN	707		
805	SRLDVCCKVQBELPPPCQSGEL-----LLG-----FFQKAYAPFLHWHITGVV	849		
708	G-----KP-EFFPPYQIKDLGAYLIGQRATDTFTMQFFHFQVAPFLHMRMTIIT	758		
850	LLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYLFLNRYFVEGAPVYFVTLGYN	909		
759	GIIFIASPIITVILSSKISVGFQDSMAFTEKSYISITHFYLDKFDVGGPPVFFVDELD	818		
910	PSSBAGMNAICSSAGCNFFSTQKIQVATEPREGSYLAIPASSWVDVDFDILT--PSSCCR	968		
819	WHRPDVQNKFTFPFGCDSTSPGNTMNYAVGTEQYLSGEMYNWIDNLEYWISRKSPCK	878		
969	LYISGPNKDKFCPSTVNSL-----NCLXNCHSITWGSVRPSVEQPHKYLPHFLNDRPNIK	1024		
879	VYVHPN--TFCSTNRNKSALDDKACRTCMDF-----	908		

Query Match	15.6%	Score 1078;	DB 2;	Length 1296;
Best Local Similarity	24.0%	Pred. No. 2.8e-92;		
Matches 321;	Conservative 237;	Mismatches 492;	Indels 288;	Gaps 41;
76	KICPLRYTGNTOACCSAKQLVLSIASITKALLTRCPACSDNPNVNLCHNTCSNQL	135		
63	EFCPHLLTGDN-KLCTSPQAGLTKQIAQHILGRPCFDNPAKLWCSEFTCSNQOD	121		
136	FINVTRVACL-GAGQLPAVVAEAF-----YQHS--FAEQSYDSCSRVRYPAAATLAVG	186		
122	FVISEMKEPIEGEFTPEYQPAEAYVNTVEYRLSTDFAEGMFSCKDVTGGQPALRV-	180		
187	TWCGYVGSALCWAQWLNFGDGTNGL-APDITPHLLPQAVGSGIOP-LNKGVARCN	244		
181	-MC---TSTPCTLTNMLEFIFGTQNDLNIPIHTFLLYDPIKTPPSRSTYMNVTGCD	236		
245	ESQGDVATSCQDCAASCPAIRPQALDSTFYLGOM-----PGSLVLIILICSVFAVVT	299		
237	KSARVGWACSTSEC--NKEEVANLIDLDGKTSQTCNVHGIACLNIFVWLAFIGSLAV	294		
300	ILLVGF-----RVAPARDKSKMVPDKGTSLSDKLSFSTHTLLGQFF	341		
295	LLCVGFVTSYDEYTNLRQTSQESPKRNRIK-----RTGAWI-----HNFMENNA	342		
342	QGWGTWASWPLTILVLSVIPVALAAGLVTELTDPVELWASAPNQARSEKAFHQHP	401		
343	RDIGMAGRNKSHFFIGCAVLIFCLPGMIYKHESNTVVDMMWSPRGRQBEVFNANF	402		
402	GPFFRTNQVILTAPNRSSRYDLSLLGPNKPSGILDLDLLELLEQLERLHLQWVSPA	461		
403	GRPQYQIMLL--SHRDFQSSGKLYG-----VPHKDIPEELFDILNAKNISTQSDG	455		
462	QRNLSQDICIAPLNPDNTSYDCINSLLQYFQNNRTLLLTANQT-----L	509		
456	-RTITLDVCYRPMGPG---YDCLINSPTNYFQCNKEHLDMKSKBETVSEDDDAFYF	510		
510	MGQTSQVMDKHDELYCANAPLTFKDGATALSCADYCAPVFPFLAIGYKGDYSAEA	569		
511	SSEATTDEWNNHMAACIDQPMQK--TKSGUSCMGTGGPSAPNN-VFGKNSNHOQANS	567		
570	LIMTFLSNYPAGDPLRLAQAKLBEAFLEENRAFQRRMAGMFQVTTAERSLEDEINRTT	629		
568	IMMTILVTQ--RTEPEIQKASLWEKEFLKFCKEYREKSPKVI-FSFWAERSITDEIENDA	624		
630	AEDLPIFATSVIVIFLYLSALGSY-----SSWRVWVDSKATGLGGVAVLGVANMA	684		
625	KDEIVTVVIALAFLIGYVTFSLGRYFCVENQWS-ILVHSR-----	664		
685	GFFSVLGRISRLVQLVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHIGRALGRVA	744		
665	GF-----	666		
745	PSMLLCSLSEACFFLGALTPMPAVRTPALTSGLAVIDLFLLOMSAFVALLSDSKRQEA	804		
667	-----TDLPAIRTFCLVAGLAVLIDVVLHCTIFLALFWDTQRELN	707		
805	SRLDVCCKVQBELPPPCQSGEL-----LLG-----FFQKAYAPFLHWHITGVV	849		
708	G-----KP-EFFPPYQIKDLGAYLIGQRATDTFTMQFFHFQVAPFLHMRMTIIT	758		
850	LLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYLFLNRYFVEGAPVYFVTLGYN	909		
759	GIIFIASPIITVILSSKISVGFQDSMAFTEKSYISITHFYLDKFDVGPVPPVFTVDELD	818		
910	PSSBAGMNAICSSAGCNFFSTQKIQVATEPREGSYLAIPASSWVDVDFDILT-PSSCCR	968		
819	WHRPDVQNKFTFPFGCDSTSPGNTMNYAVGTEQYLSGEMYNWIDNLEYISRSKPCCK	878		
969	LYISGPNKDKFCPSTVNSL-----NCLXNCHSITWGSVRPSVEQPHKYLPHFLNDRPNIK	1024		
879	VYVHPN--TFCSTNRNKSALDDKACRTCMDF-----	908		

Db 24 PGRPAGGRRRTTGGRLRAAADPRD-----YLHRPSYCDAAFALEQI--- 65
Qy 285 LVLIILILCSLSEAVVITLLVGRVAPADKSKRWDPK-----RGTSLSDKLSPSTHTLL 337
Db 66 -----SKGKATGRKAPLWRAKFORLLFKLGYIQKNC 98
Qy 338 GQFFOGWGTWASWPLTILVSVIPVVALAGLVEITLTPDVELSNAPNSQARSEKAFH 397
Db 99 GKF-----LVGGLIFGAFVGLKAAANLETNVEELWVEVGGVRSRELYT 143
Qy 398 DQHPGPPFRFN-QVILTAPNRSSRYDSSLGPKNFSGILDLLELLE--LQERLRLH 454
Db 144 RQKIGEEAMFNQLMIQPKKEG-----ANVLTEALLQHLDSALQASRVHV 190
Qy 455 QWSPEAQRNLSQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 191 YMYN-----RQWKLCHLYKSGBELITETGMDQIIEVLYPCLIIITPLDCFEGWAKLQSGTA 246
Qy 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFDPLEFLEBELKKNYQVDSWEEMLNKAIEVGHYMDRPLCLNADPDC 304
Qy 528 -----APL-----TFKQGT-----ALALSQWA 544
Db 305 PATAPNKSTKPLDMALVNLGCHGSLRKYMWQBELIVGGTVKNSGTGLVSAHALQTMF 364
Qy 545 DYCAPVFPFLAIGYKGYDSEAEALIMTFSLNYPAGDPRLQAQAKLWEEAFLEMRAPQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE---DKAAAILLEAQWRTVEVHVQSV 411
Qy 605 RMAGHFQVTFARSLDEINRTAEDLPITATSYIVIFLISALGSYSWSRVWVDS 664
Db 412 AQNSTQKVLSTFT--TTLDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
Qy 665 KATLGIGGVAVLGVAYMAAGPFSYLGIRSLVILQVVPFLVLSVADNIFIVLEYORL 724
Db 466 QGAVGLAGVLLVALSVAGLGLCSLIGISFNATTOVLPFLALGVGVDDVFLAHAFST 525
Qy 725 PRPGPBPVHIGRALGRVAPSMLLCSLSEACFFLIGALTMPAVNTFALTSGLVILDF 784
Db 526 GQNKRIPTEDRTGECRLKRTGASVALTSISNVTAFMAALIPALRAFSLQAAVVVVFNF 585
Qy 785 LLOMSAFVALLSDSKQESRLDVCCC-----VKQEL-----PPG 822
Db 586 AMVLLIFPAILSDMLYRDRRLDIFCCFTSPCVSRVIVQEPQAYTDTHTNRYSPPPY 645
Qy 823 QGEG----- 827
Db 646 SSHSFAHETQITMQSTVOLRTEYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGPFQKAYAPFLHWTIRGVVLLFLALFGVSL 861
Db 706 SSTRDLLSQFSSSLHCLPEPCTKWTLSSPAETHYAPFLKPKAKVWVIFLFLGLLGVSL 765
Qy 862 YSKCHISVGLDQELALPKDSYLDYFLFLNRYEVGAPVYFVTLGVNFSSEAGMNAICS 921
Db 766 YGTTVRVDRGLDLDIVPRETREYDFTAAQKYSF-----YNM----- 803
Qy 922 SAGCNFSFTQKIQYAT-----EPPEQSYLAIPASS-----WVDDFIDL----- 961
Db 804 -----YIVTQKADYFNIQHLLYDLHRSFSNVKYVWLEENKQLPKMHLHYFRDLWLOGLQD 857
Qy 962 -----TPSS-----CCRLVISGNKDKFCPSTVNSLNCNCKMSITMGS 1000
Db 858 AFDSDETQKIMPNYKNGSDGVLYAKLLVQTSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQFHYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLTSQDQ 1043
Db 914 INPSA--FYIYLTAVNSNDPVYAASQANTIRPHRPEWHDK--ADYMPETRLRIPAAEP 968
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA--DLRKVPQDTPAFEPFVYITNIFY 1102
Db 969 IEYAQFPPLNGLRDTSDFEABEKVKTICSNTYSLGLSSYPNG-----YPF-----LFW 1018

Qy 1103 EQYLITLPEGLFMLSCLVPTRAVSCLLGLDLRSGLLNLLSIWMLIVDTVGFMAWDIS 1162
Db 1019 EQYIGLRHMLLISVVLVACTFLVCAVFLNPTAGII--VWVLATMTVELFGMGLIGIK 1077
Qy 1163 YNAVSLINLVSAGMGSVEFVSHITRSF--AISTKPTWLERAKGATISMGSVAVFAGVAMTN 1220
Db 1078 LSAVPPVILLIASGVIGVEFTVHVALAFLTAIGDKN---RRAVLALAHMFAPVLDG-AVST 1133
Qy 1221 LPGILVLGLAKAQILQIPFFRLNLLITLLGLLHGLVPLPVLISVVG--PDVNPALALEQ- 1277
Db 1134 LLGLVMLAGSEFDFIVRYFAVLAILTILGLVNLGLVLLPVLISFFGYPEVSPANGNLRL 1193
Qy 1278 --KEABEAAVAVMVASCPNHPKSRVSTADNIIYVNHSPGSIKG 1317
Db 1194 PTPSPPEPPSVVRPAMPFGHTH--SGSDSDSDSEYSSQTTVSG 1233

RESULT 7
US-08-656-055-19
; Sequence 19, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
Qy 225 PGQVAGSGIQPLNEGVARCNESQGDVATCSCODCAASCPATARPALLSTFVLGMPGS 284
Db 24 PGRPAGGRRRTTGGRLRAAADPRD-----YLHRPSYCDAAFALEQI--- 65
Qy 285 LVLIILILCSLSEAVVITLLVGRVAPADKSKRWDPK-----KGTSLSDKLSPSTHTLL 337

Db 66 -----SKGATGRKAPLWRAKFORLLFKLGCYIQKNC 98
Qy 338 GQFQGWGTWASWPLTILVSPVVALAAGLVFTELTDPVELWASPNQARSEKAPH 397
Db 99 GKF-----LVVGLLIFGAPAVGLKAANLETVBELWVEVGGVRSRELYT 143
Qy 398 DQHFQFPFRTN-QVILTAPNRSSRYRDSLLGPKNFSGILDLDLLELLE--LQERLRLH 454
Db 144 RQKIGSEAFNPQMTQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190
Qy 455 QVMSPEAQRNISLQDICYAP-----LNPDTSLYDCINCNSLLQYFQNNRTLLLLTA 505
Db 191 YMN-----RQMKLEHLCYKSGELITETGYMDQIIYVLPCLITPLDCFWEGAKLQSGTA 246
Qy 506 NOTLMGQ-----TSQVMDKHPLY-----CAN-----527
Db 247 --YLLGKPLRWTFNFDPLFLEBELKKINYQVDSWEEMLNKAEGVGHGMDRPCLPNADPDC 304
Qy 528 -----APL-----TFKDT-----ALALSCHA 544
Db 305 PATAPNKNSTKPLDMALVNLGCHGSLRKYMHQWELIVGGTVKNSGTGLVSAHALQTMF 364
Qy 545 DYCAPVPEPLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAALWEEAFLEEMRAFQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINNE-----DKAAALEAWORTTVEVVHQSV 411
Qy 605 RRMAGNFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSYSMSRWVMS 664
Db 412 AQNSTQKVLSTFT--TTLDLILKSPSDSVIRVASGYLLMLAVACLTM---LRWD--CSKS 465
Qy 665 KATLGLGVAVVLGVAWMAANGFSSYLGRSSVLIOVFPFLVLSVGADNIFPVLVEYQRL 724
Db 466 OGAVGLAGVLLVALSVAAGLGCSTLIGSFNAATTQVLPFLALGVGDVDFLLAHAFSET 525
Qy 725 PRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPAVRTTALTSGLAVILDP 784
Db 526 GQNKRIFFEDRTGECLEKRTGASVALTSISNTVAFFMAALIPALRAFSLQAAVVVNF 585
Qy 785 LQMSAFVALLSDSKQBSRLDVCCC-----VKPQEL-----PPPG 822
Db 586 AMVLLIFPAILSMWLYRREDRLDIFCCFTSPCVSRVIOVEPQAYTDTHDTRYSPPPY 645
Qy 823 QEGEL-----827
Db 646 SSHSFAHETQITWQSTVOLRTEYDPTHVYVYTAEPSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGFFOKAVAPLHWTIRGVLLLLFLALFGVSL 861
Db 706 SSTRDLSQFSSSLHLEPPCTKWTLSFAEKHYAPFLKPKAKVWVIFLGLLGVSL 765
Qy 862 YSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTLGYNPSSEAGNNAICS 921
Db 766 YGTVTRVGLDLTDIVPRETRYEDFIAQPKYFSP-----YNN-----803
Qy 922 SAGCNFSFTQKIQVAT-----BPPEQSYLAIAPASS-----WVDDFDIDLW-----961
Db 804 -----YIVTQADYFNIGHLLYDLHRSFNKYYVMLEENKQPKMWLHYPRDMLQGLQD 857
Qy 962 -----TPSS-----CCRLYISGPNKDKPCSTVNSLNCNKMISITWGS 1000
Db 858 AFDSWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVADAGI 913
Qy 1001 VRPSVEQFHXYL-PWFLND-----RPNKCPKGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYIYLTAVSNDPVAYASQANIRPHRPFWHDK---ADYMPETRLRPAEP 968
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPFVPPYTTINVFY 1102
Db 969 IEYAQFPFPLNGLRDTSDFVEAIEKVRTICSNTYSLGLSSYPNG-----YPF---LFW 1018
Qy 1103 EOYLTLPEGLPMLSLCVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWDIS 1162
Db 1019 EOYIGLRHWLLFISVVLACTPLVCVAFLLNPWTAGII--VMVLALMTVELFGMGLIGK 1077

Qy 1163 YNAVSLINLVSAVMSVEFVSHITRSF--AISTKPTWLBRAKEATISMGSAVPAVAMTN 1220
Db 1078 LSAVPVILLIASVGIVGEFTVHVALAFLTAIGDN---RRAVLAEHMFAPVLDG-AVST 1133
Qy 1221 LFGILVGLAKAQLQIOPFRNLNLLITLGLHGLVFLPVILSYVG--PDVNPALLAEQ- 1277
Db 1134 LLGLVMLAGSBDFIVRYFFAVLAITLITGLVNLGLVLLPVLSSFFGYPYEVSPANGLNRL 1193
Qy 1278 --KRAEEAAVAAVMVASCPNHPRSVSTADNIYVNHSPESGSIK 1317
Db 1194 PTPSEPPEPPSVVRPAMPPOGHTH--SGSDSDSEYSQTTVSG 1233

RESULT 8

US-08-954-668-19
; Sequence 19, Application US/08954668
; Patent No. 6172200
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,668
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-668-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;

Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PQAVGSGIQPLNEGVARNESQGDVATCSCQDCAASCPAIPARQALDSTFYLGQMPGS 284
Db 24 PGRPAGGRRRTTGLRRAAPDRD-----YLHRRSYCDAAAFALQI---65
Qy 285 LVLIILICSVAVVTILLVGRFVAPARDKSKMVDPK-----KGTSLSDKLFSFTHLL 337
Db 66 -----SKGATGRKAPLWRAKFORLLFKLGCYIQKNC 98
Qy 338 GQFQGWGTWASWPLTILVSPVVALAAGLVFTELTDPVELWASPNQARSEKAPH 397
Db 99 GKF-----LVVGLLIFGAPAVGLKAANLETVBELWVEVGGVRSRELYT 143
Qy 398 DQHFQFPFRTN-QVILTAPNRSSRYRDSLLGPKNFSGILDLDLLELLE--LQERLRLH 454

Db	144	RQKIGEEAMFNPQMIQTPEKEG-----ANVLTTEALLQHLDSALQASRVHV	190
Qy	455	QVWSPQARNISLQDICYAP-----LNPDNSTLYDCCINSLLQVFNRRFLLLLTA	505
Db	191	YMYN-----RQWLEHLCYSGELITETGYMQIIEYLYPCLIIITPLDCFWEAKLQSGTA	246
Qy	506	NOTLMGO-----TSQVDWKDHFY-----CAN-----	527
Db	247	--YLLGKPLRWTFNDFPLELEELKKINYQVDSWEEMLNKAEVGHGYMDRPLCPADPDC	304
Qy	528	-----APL-----TFKQGT-----ALALSCMA	544
Db	305	PATAPNKNSTKPLDMALVINGCGHLSRKRYMHWQBELIVGTVMNSTGKLVSAHALQTMF	364
Qy	545	DYGAPVFPFLAIGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEPLSEEMRAFQ	604
Db	365	QUMTPKQMY---EHPKGYEY-----VSHINWNE-----DKAAAILLEAQRTYVEVHQSV	411
Qy	605	RMAGMFQVTFTAERSLEDEINRTTAEDLPFIPATSYIVIFLYISIALGYSYSSWSRVMDVS	664
Db	412	AQNSTQKVLSTFT--TTLDILKSPFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS	465
Qy	665	KATLGLGGVAVVLGVAAMAAGFFSYLGRSSVLIVQVPPFLVSVGADNIRIFVLEYQRL	724
Db	466	QAGVLAGVLLVALSVAAGLGLCSLIGSFNAATTQVLPFLALGVGVDDVFLLAHAFSET	525
Qy	785	LLQMSAFVALLSLDKROBASRLDVCCC-----VKPQEL-----PPGQEL	822
Db	586	AMVLLIFPAILSDMLYRREDRLDIFCCFTSPCVSRVIOVBPQAYTDDHNDTRYSPPPPY	645
Qy	823	QGBGL-----	827
Db	646	SSHSPAHETQITWQSTVQLRTEYDPHTHVYTTAEPRSEISVQPVTVTQDTLSCQSPST	705
Qy	828	-----LLGFFQKAYAPFLLHMTIRGVVLLFLALFGVSL	861
Db	706	SSTRDILLSQFSSSLHCLBPPCTKTWLTSSFAEKHYAPFLLKPKAKVWVIFLGLLGVSL	765
Qy	862	YSMCHISVGLDQBSLALPKXSYLDYFLFNRYEYVGPVYVFTTLGYNFSSSEAGMAICS	921
Db	766	YGTTRVRDGLDLDIVPRETREYDFIAAQPKYFSF-----YNM-----	803
Qy	922	SACNNFSTQKTOYAT-----EPEQSYLAIPASS-----WVDDFIDWL-----	961
Db	804	-----YIVTQKADYPNIOHLLYDLHRSFSNVKYVMLEENKQLPKWHLHYFRDNLQLOD	857
Qy	962	-----TPSS-----CCRLYISGNKDKFCPSTVNSLNCNKMCSITMGS	1000
Db	858	AFDSDWETGKIMPNYKNGSDGGLVAYKLLVQTSRDK--PIDISQLT--KORLVDADGI	913
Qy	1001	VRPSVEQFHXYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLTSQDQ	1043
Db	914	INPSA--FYILTAWVSNDFVAVAAISOANIRPHRPWVHDK---ADYMPETRLRIPAAEP	968
Qy	1044	VLASRFWAYHKPLKNSODYTEALRAARELANITA--DLRKVPGTDPAPFEPVYITNVFY	1102
Db	969	IEXAQFPFYNGLRDTSDFVAIEKVRTICSNYTSGLSSYPNG-----YFP-----LFW	1018
Qy	1103	EQYLTILPEGLFMLSCLVPTFAVSCLLLGLDLRSLGNLLLSIVMLIVDTVGFMALWDIS	1162
Db	1019	EQYGLRHWLLLFISVVLACTFLVCVAFLLNPWTAGII--VMVALMTVELFGMMGLIGIK	1077
Qy	1163	YNVAVSLINLNSAVGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1078	LSAVPVVILIASVGTGVEFTVHVVALAFLTAIGDKN---RRAVLALEHMFAPVLDG-AVST	1133
Qy	1221	LPGLIIVGLAKAQLIQIFPERLNLITLLGLLHGLVFLPVILSYVG--PDVNPALALEQ-	1277

Db	1134	LLGVLMAGSEPDFIVRYFFAVLAILTILGLVINGLVLPLVLSFFGFPYPEVSPANGLNRL	1193
Qy	1278	--KRAEEAANAAMVASCNPHPSRVSTADNIYVNVHSPGSGIKG	1317
Db	1194	PTSPPEPPSVVRFAMPPPGHTH--SGSDSDSDSYSSQTTVSG	1233
RESULT 9			
US-09-268-140-5			
; Sequence 5, Application US/09268140			
; Patent NO. 6268176			
; GENERAL INFORMATION:			
; APPLICANT: Gemmill, Robert M.			
; APPLICANT: Drabkin, Harry A.			
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED			
; FILE REFERENCE: 93445-00004			
; CURRENT APPLICATION NUMBER: US/09/268,140			
; CURRENT FILING DATE: 2000-03-12			
; PRIOR APPLICATION NUMBER: US 60/077,723			
; PRIOR FILING DATE: 1998-03-12			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: Patentin Ver. 2.0			
; SEQ ID NO 5			
; LENGTH: 1447			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-268-140-5			
Query Match 9.4%; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	RQAGVSGIGIPLNEGVARCNESQGDVATCSCQDCAASCPAIAFPALDSTFYLGMPGS	284
Db	24	PERPAGGRRRTTGLRRAAAPDRD-----YLHRSYCDAAFALEQI---	65
Qy	285	LVLIIILCSVFVAVTILLVGRFVAPARDKSKRWDPK-----KGTSLSDKLSFSTHILL	337
Db	66	-----SKGATGRKAPLWLRAKQRLFLKLCGYIQKNC	98
Qy	338	QGFPGWGTWASWPLTTLIVLSPVVALAAGLVFTELTDPVELWSAPNSQARSEKAPH	397
Db	99	GKF-----LVGLLIFGAPAVGLKAANLETNVEELWVEVGGVSRSELNTY	143
Qy	398	DOHFQFPFRTN--QVILTAPNRSRYSDSLLGPKNFSGLDLDLLELE--LQERLRL	454
Db	144	ROKIGEEAMFNPQMIQTPEKEG-----ANVLTTEALLQHLDSALQASRVHV	190
Qy	455	QVWSPQARNISLQDICYAP-----LNPDNSTLYDCCINSLLQVFNRRFLLLLTA	505
Db	191	YMYN-----RQWLEHLCYSGELITETGYMQIIEYLYPCLIIITPLDCFWEAKLQSGTA	246
Qy	506	NOTLMGO-----TSQVDWKDHFY-----CAN-----	527
Db	247	--YLLGKPLRWTFNDFPLELEELKKINYQVDSWEEMLNKAEVGHGYMDRPLCPADPDC	304
Qy	528	-----APL-----TFKQGT-----ALALSCMA	544
Db	305	PATAPNKNSTKPLDMALVINGCGHLSRKRYMHWQBELIVGTVMNSTGKLVSAHALQTMF	364
Qy	545	DYGAPVFPFLAIGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEPLSEEMRAFQ	604
Db	365	QUMTPKQMY---EHPKGYEY-----VSHINWNE-----DKAAAILLEAQRTYVEVHQSV	411
Qy	605	RMAGMFQVTFTAERSLEDEINRTTAEDLPFIPATSYIVIFLYISIALGYSYSSWSRVMDVS	664
Db	412	AQNSTQKVLSTFT--TTLDILKSPFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS	465
Qy	665	KATLGLGGVAVVLGVAAMAAGFFSYLGRSSVLIVQVPPFLVSVGADNIRIFVLEYQRL	724
Db	466	QAGVLAGVLLVALSVAAGLGLCSLIGSFNAATTQVLPFLALGVGVDDVFLLAHAFSET	525
Qy	725	PRRPGEPREVIHGRALGRVAPSMLLCSLSEAI-CFFLGLALTTPMPAVRTFALTSGLAVIDF	784

Db 526 GQNKRIFFEDRTGECUKRTGASVALTSISNVTAFMAALIPALRAFSLQAAVVVFN 585
QY 785 LLOMSAFVALLSDSKRQBSRLDVCC-----VKQEL-----PPPG 822
Db 586 AMVLLIFPAILMDLYRDRRLDIFCCFTSPCVSRVQVPEQAYTDHNTRYSPPPY 645
QY 823 QSEGL----- 827
Db 646 SSHSFAHETITWQSTVQLRTEYDPRHVTYHTAEPREISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGFFOKAYAPPELLHWTIRGVILLFLALFGVSL 861
Db 706 SSTRDLLSOPSSSLHLEPPCKWTLSFAEKHYAPFLPKAKVWVIFLGLLGVSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTTLGNFSSBAGMAICS 921
Db 766 YGTRVRDGLDITDIPRETREYDFAAQFKYF-----YNN----- 803
QY 922 SAGCNPFSTQKIYAT-----EPPEOSYLAIPASS-----WVDDFDWL----- 961
Db 804 -----YIVTORADYENIOHLLYDLHRSFSNVKYVMLEENKQLPKMWLHYFRDMLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDKCPSVTNSLNCCLKMCSITMGS 1000
Db 858 AFDSWETGKIMPNYKNGSDGVLAYKLLVOTGSRDK--PIDISQLT--KORLVDADGI 913
QY 1001 VRPSVEQFKYL--PWFIND-----RPNKCPKGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYILTAVNSNDPVAYASQANIRPHREWHDK--ADYMPETRLRIPAAEP 968
QY 1044 VLASRPMATHKPKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPFVFFYTTNVFY 1102
Db 969 IEYAQFPFYINGLRDTSDFEAEIKVRTICSNVYSLGSSYPNG-----YFF-----LFW 1018
QY 1103 EOYLITLPEGLWLSICLPTFAVSCLLGLDLSGLNLLSIVMLVDTVCFMALWDIS 1162
Db 1019 EOYIGLRHLLFISVLACTFLVCAVFLNFWTAGII--VMVALMTVELFGMGLIGIK 1077
QY 1163 YNAVSLINLVSAGMSVEFVSHITSF--AISTKPTWLERAKEATISMGSAFAGVAMTN 1220
Db 1078 LSAVPVILIASGVGEVETVHVALAFLTAIGDKN---RRVLALEHMPAPVLDG-AVST 1133
QY 1221 LPIGLVLGAQALQIOPFFRLMLLITLGLLHGLVFLPVILSYVG--PDVNPALALIEQ- 1277
Db 1134 LILGVLMLAGSEDFIVRYFPAVLAILITLGLVNLGLVLLPVLLSPFGPYEVSFANGNRL 1193
QY 1278 --KRAEEAAMVAVASCPNHPRSVSTADNIYNHSPGSIKG 1317
Db 1194 PTPSPPPPPSVVRFAFPPGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 10

US-08-918-658-19

; Sequence 19, Application US/08918658

; Patent No. 6429354

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GOODRICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/918,658

; FILING DATE: 22-Aug-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/656,055

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/540,406

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1447 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

; US-08-918-658-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;

Best Local Similarity 22.0%; Pred. No. 1.1e-51;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQQAVSGIGTQPLNEGVARCNESQGDVATCSCQDCAASCPAIAARPOALDSTFVLGQMPGS 284

Db 24 PGRPAGGRRRTTGLRAAAPDRD-----YLHRSYCDAAFALEQI--- 65

QY 285 LVLIILICSVFAVVTILLVGFVAPARDKSKWVDPK-----KGTSLSDKLSFSFHTLL 337

Db 66 -----SKGKATGRKAPLWLRKAFQRLFLKGCYIQKNC 98

QY 338 GQFFQGWGTWASWPLTILVLVSVIPVVALAAGLVTELTTPVELMSAPNSQARSEKAFH 397

Db 99 GKPF-----LVGLLIFGAPAVGLKAALETNTVEELNVEVGGVRSRELNTY 143

QY 398 DQHGFFPFRTN-QVILTAPNRSSRYSDLSLLGPKNFSGILDLDLLELE--LOERLRHL 454

Db 144 RQKIGEEAFWPLMIQTPKSEG-----ANVLTTEALLQHLDSALQASRVAV 190

QY 455 QVMSPEAQRNLSLDICYP-----LNPDNLSYDCCINSLLQVFNRTLLLLTA 505

Db 191 YMYN-----RQWKLHLCYKSGELITETGYMDQIIEYLYPCLITPLDCFWEGAKLQSGTA 246

QY 506 NQTLMGQ-----TSQVDWKDHFY-----CAN----- 527

Db 247 --YLLGKPLRWTFNPDPLELEELKKINYQVDSWEEMLNKAERVGHGYMDRPCINPADPDC 304

QY 528 -----APL-----TFKQGT-----ALALSCWA 544

Db 305 PATAPNKNSTYPLDMALVINGCHLSRKYMHVQBELIVGGTVKNSTKGLVSAHALQTMF 364

QY 545 DYGAVPVFPFLAIGGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFQ 604

Db 365 QLMTPKQWY---EHPFKGYEY-----VSHINWNE---DKAAAILEAMQRTVEVVHQS 411

QY 605 RRMAGMFQVTTAERLSLEDEINRTAEDLPITPATSYIVIFLYISALGSYSWSRWVDS 664

Db 412 AQNSTQKVLSTPT--TTLDILKSPSDSVIRVASGYLLMLAYACITM---LRWD--CSKS 465

QY 665 KATLIGGVAVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSVGADNIFLVEYQRL 724

Db 466 QGAVGLAGVLIVALSAGLGLCSLIGISFNAAATQVLPFLALGVGVDDVFLAAHAFSE 525

QY 725 PRPGEPREHVHGRALGRVAPSMILCSLSEIACFFLGTALTPMPAVRTFALTSGVLIDF 784

Db	646	SSHFAHETQITMSTVQLRTEYDPHTHYTTAEPRSEISVQPVTVTQDTLSCQSPST	705
Qy	828	-----LLGFFQKAYAPFLLHWITRGVVLFLFALFGVSL	861
Db	706	SSTRDLLSQFSDSSHLCLPEPCTKWTLSFPAEKHYAPFLPKAKVWVIFLFLGLLGVSL	765
Qy	862	YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFTTLGYNFSSEAGHNAICS	921
Db	766	YGTTRVRDGLDLDIVPRETREYDFIAAQFKYFSF-----YNN-----	803
Qy	922	SAGCNPFSTQKIQYAT-----EPQSYLAIPASS-----WDDDFIDWL-----	961
Db	804	-----YIVTQKADYPNIQHLLDYLHRSFNVKYVWLEENKQLPKMWLHYFRDWMLOGLQD	857
Qy	962	-----TPSS-----CCRLYISGPNKDKFCPSTVNSLNCNKCMSTMG	1000
Db	858	AFDSDWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI	913
Qy	1001	VRPSVEQPHYL--PWFLND-----RPNKCPKGLAAY--STSUNLTSQDQ	1043
Db	914	INPSA--FYIYLTAWVSNDDPVAYAASQANIRPHREPWHDK---ADYMPETRLRIPAAEP	968
Qy	1044	VLASRFMAHYKPKNSQDYTEALRAARELAANITA--DLRKVPQTDPAFEVFPYTTITNVFY	1102
Db	969	IYEAQFPFYLNGLRDTSDFVEAIEKVRTICSNYTSIGLSSYPNG-----YFP-----LFW	1018
Qy	1103	EQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLLNLLSIVMLVDTVVGFMALWDIS	1162
Db	1019	EQYIGLRHMLLFISVVLACTFLVCAPFLNPNWTAGII--VMVLALMTVELFGMGLIGIK	1077
Qy	1163	YNAVSLINLVSAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMSAVFAGVAMTN	1220
Db	1078	LSAAPPVILLIASVGIGVEFTVHALAFLTAIGDN---RRVLALEHMFAPVLDG--AVST	1133
Qy	1221	LPGLVLGLAKAQIIGIFFRNLMLITLGLHGLVFLFVILSVYG--PDVNPALALEQ-	1277
Db	1134	LQVLMWLAGSEDFIVRYFAVLAIIITILGVNLGLVLLPVLVSFFGYPPEVSPANGLNRL	1193
Qy	1278	--KRAEEAANAAMVASCNPHGRVSTADNIYVNHSPGSIKG	1317
Db	1194	PTSPSPPPSVVRFAFMPGHTH--SGSDSDSEYSSQTTVSG	1233
RESULT 13			
US-09-807-007-6			
; Sequence 6, Application US/09807007			
; Patent No. 6881833			
; GENERAL INFORMATION:			
; APPLICANT: ZAPHIROPOULOS, Peter et al.			
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY			
; FILE REFERENCE: 2921-0130P			
; CURRENT APPLICATION NUMBER: US/09/807,007			
; CURRENT FILING DATE: 2001-04-06			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 6			
; LENGTH: 1447			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-807-007-6			
Query Match 9.48; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVSGGIOPLENGVARCNESQDDVATCSODCAASCFAIARPOALDSTFYLGQMPGS	284
Db	24	PGRPAGGRRRTTGLRAAAPRED-----YLHRPSYCDAAFALEQI---	65
Qy	285	LVLIIILCSFVAVVTIILLVGRVAPARDKSMVDPK-----KGTSLSKLSFSTHTLL	337
Db	66	-----SKGATGRKAPLMLRAKFORLLFKLGCYQKNC	98


```

Db      1078  LSAPVWIIASVGVETVHVALAFLTAIDGN---RRVLALEHMFAPVLDG-AVST 1133
Qy      1221  LFGILVLGLAKAQLIQIFFRMLNLITLLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277
Db      1134  LLGVLMAGSEPDFVIRYFFAVLAITLILGVNLGLVLLPVLSPFGPYEVSFANGNLRL 1193
Qy      1278  --KRAEAVAAMVWASCPNHPRSRVSTADNIYNHSPGSIKG 1317
Db      1194  PTPSPPPSVVRFPMPGHTH--SGSDSSDSEYSSQTTVSG 1233

```

RESULT 14

```

US-09-754-032-19
; Sequence 19, Application US/09754032
; Patent No. 6921646
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
;              GOODRICH, LISA V
;              JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,032
; FILING DATE: 03-Jan-2001
; CLASSIFICATION DATA:
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-Oct-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
;
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

```

```

Query Match      9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

```

```

Qy      225  PQQAVSGIOPLNEGVARNESQGDVATCSQDCAACPALARPQALDSTFYLGQMPGS 284
Db      24   PGRPAGGRRRTTGLRRAPDRD-----YLHRSYCDAAFALEQI--- 65
Qy      285  LVLIILILCSFAVWVITLLVGRFVAPARDKSKMVDPK-----KGTSLSDKLSEFSTHTLL 337
Db      66   -----SKGKATGRKAPLWLRAKFORLLFKLGCYIQNC 98
Qy      338  GQFFQCGTGWASWPLTILVLSVIPVVALAAGLVTELTTPDVELMSAPNQARSEKAPH 397
Db      99   GKF-----LVVGLLIFGAPVGLKAAANLETNVEELWVEVGGVRSRELNYT 143

```

```

Qy      398  DQHFGPFFRTN-QVILTAPNRSSRYSDSILLGPKNPSGILDLDLLELE--LQERLRL 454
Db      144  RQKIGEEAMFNQMIQTPKEBG-----ANVLTTEALLQHLDLSALQASRVHV 190
Qy      455  QWMSPEAQRNLSLDICVAP-----LNDPNTSLYDCCINSLLQYQNNRTLLLLTA 505
Db      191  YMYN----ROWKLEHLCKYSGBELITETGYMDQIIBLYPCLITITPLDCFEWAKLQSGTA 246
Qy      506  NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db      247  --YLLGKPLRWNTNFDPLFLELKKINYQVDSWEMLNKAHVGHGMDRPCLNADPDC 304
Qy      528  -----APL-----TFKQCT-----ALALSCMA 544
Db      305  PATAPNKNSTKELDMALVNGGCHGLSRKYMHWQBELIVGGTVKNSGKLVSAHALQTMP 364
Qy      545  DYGAVPFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEENRAFO 604
Db      365  QLMTEPKQMY--EHFKGYEY-----VSHINWNE---DKAAAILEAMQRTYVEVVHQS 411
Qy      605  RRMAGMFQVTTAERSLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSWSRVMVDS 664
Db      412  AQNSTQKVLSPFT--TLDDILKSPDSVSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
Qy      665  KATLGLGGVAVVVLGAVMAAGFFSYLGIRSSILVILQVVPFLVLSVGADNIPFLVLEYQRL 724
Db      466  QGAVGLAGVLLVALSVAAGLGLCLIGISFNAATTVQVLPFLALGVGVDDVFLAHAFST 525
Qy      725  PRPCEPREVHIGRALGRVAPSMMLCSLSEACFFLGLALTPMPAVRTFALTSLGLVILDF 784
Db      526  QONKRIPFEDRTGECLKRTGASVALTISNVTAFFMAALIPIPALRAFSLQAAVVVVF 585
Qy      785  LLQMSAFVALLSLDSKROEASRLDVCCC-----VKQBEL-----PPG 822
Db      586  AMVLLIFAILMSDLYRREDRRDLIFCCFTSPCVSRVIOVEPOAVTTHDNTRYSPPPY 645
Qy      823  QGEGL----- 827
Db      646  SSHSPAHEQTIMQSTVQLRTEYDPTHVYVYTAEPSEISVQPVTVTQDTLSCQSPST 705
Qy      828  -----LLGFQKAYAPFLHWHITRGVLLFLALFQVSL 861
Db      706  SSTRDLLSQFSDSSLHCLPEPCTKWTLSSFAEKHYAPLLPKAKVWVIFLGLLGVSL 765
Qy      862  YSMCHISVGLQDELALPKDSYLLDYFLNRYFFEGAPVYFVTTLGYNPFSSAGHNAICS 921
Db      766  YGTRVRDGLDLDIVPRETREYDFIAAQKYFSF-----YNN----- 803
Qy      922  SAGCNPFSTQKIQYAT-----EFPEQSYLAIAPASS-----WVDDDFIDWL----- 961
Db      804  -----YIVTQKADYPNIQHLLYDLHRSFNVKYYVWLEENKQLPKMWLHYFRDMLQGLQD 857
Qy      962  -----TPGS-----CCRLYISGPNKDKPCPSTVNSLNCNCKMSITMGS 1000
Db      858  APDSDWETGKIMPNNYKNGSDGVLYAKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
Qy      1001  VRPSVEQFKHYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLTSDGQ 1043
Db      914  INPSA--FYIYLTAWVNSNDPVAYAASQANIRPHRPFWHDK---ADYMPETRLRIPAAAP 968
Qy      1044  VLASRFMAHYKPKNSQDYTEALRAARELANITA-DLRKVPGTDPAPFVPPYTTITNVPY 1102
Db      969  IEYAQFPYINGLRDTSDFVEAIEKVRTICSNYISLGLSSYPNG-----YFP-----LFW 1018
Qy      1103  EGYTILPEGLPMLSLCLVPTFAVSCLLGLDLSGLNLLSIVMLIVDTVGFMALWDIS 1162
Db      1019  EGYTLRHWLLLFISVVLACTFLVCAVFLNPWTAGII-VMVLAJMTVELFGMGLIGIK 1077
Qy      1163  YNAVSLINLVSAGMSVEFVSHITRSP--ALSTKPTWLERAKEATISGSAVFAVAMTN 1220
Db      1078  LSAVPWVILIASVGVETVHVALAFLTAIDGN---RRVLALEHMFAPVLDG-AVST 1133

```

```

QY 1221 LFGILVLGAKAQIQLIOPFFRLNLLITLLGLLHGLVFLPVLSYVG--PDVNPALALEQ- 1277
DB 1134 LGLVLMAGSEDFIVRYFAVLAAILITLGLVNLGLVLLPVLLSPFGPYEVPSPANGLNRL 1193

QY 1278 --KRAEAAVAVMVASCPNHPSPVSTADNIYVNHSPSGSIKG 1317
DB 1194 PTPSPPPPSVVRFAAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 15
US-08-916-140-19
; Sequence 19, Application US/08916140
; Patent No. 6946257
; GENERAL INFORMATION:
; APPLICANT: Scott Mathew P.
; APPLICANT: Goodrich, Lisa V.
; APPLICANT: Johnson, Ronald L.
; APPLICANT: Epstein, Ervin Jr.
; TITLE OF INVENTION: PATCHED GENES AND USES RELATED THERETO
; FILE REFERENCE: CIBT-P04-203
; CURRENT APPLICATION NUMBER: US/08/916,140
; CURRENT FILING DATE: 1997-08-21
; PRIOR APPLICATION NUMBER: US 08/656,055
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: US 08/540,406
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: US 08/319,745
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-08-916-140-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGQAVSGTQPLNEGVARCNESQGDVATCSQDCAASCPAARPOALDSTFVLGMPQS 284
DB 24 PGRPAGGRRRTGTLRAAAPDRD-----YLHRPSYCDAAFALEQI--- 65

QY 285 LVLIILCSFVAVVITLLVGFVAPARDKSMVDPK-----KGTSLSDKLSFSHTLL 337
DB 66 -----SKGATGRKAPLWLRKAFORLLFKLGCYIQKNC 98

QY 338 GQFFQGGTGWASWPLTILVSVVVAAGLVFTTELTDPPVELWSAPNSQARSEKAFH 397
DB 99 GKP-----LVUGLLIFGAFVGLKAANLETNVEELMVEVGGVRSRELNT 143

QY 398 DQHPGPFRTN-QVILPAPNRSYRSDLLGKPNFSGILDLLLELLE--LQERLRL 454
DB 144 RQKIGBEAMFNPMQITQKEG-----ANVLTTEALLQHLDSALQASRVHV 190

QY 455 QVNSPEAQRNLSQDICYAP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLTA 505
DB 191 YMYN-----RQWLEHLCKYSGELITETGYMDQIIEYLYPLIITPLDCFWEKALQSGYA 246

QY 506 NQTLWGO-----TSQVDMKDHFLY-----CAN----- 527
DB 247 --YLLGKPLPWTNFDPLFLEELKKNYQVDSWEEMLNKAENVGHGYMDRPLNADPDC 304

QY 528 -----APL-----TFKDG-----ALALSCMA 544
DB 305 PATAPNKNSTKFLDMALVNLGGCHLSRKYMHQOEELIVGTVKNSTGKLVSAHALQTMF 364

QY 545 DYGAIPVFPFLAIGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMAFQ 604
DB 365 QLMTPKQMY-----EHFKGYEY-----VSHINWNE-----DKAAAILLEAQRITYVEVHQSV 411

```

```

QY 605 RRMAGMPQVTFTAERSLEDEINRTTAEDLPIRATSVIVIFLYISLALGSYSSSRVMVDS 664
DB 412 AQNSTQKVLSTFT--TTDDILKFSQSVIRVAGSGLMLLAYACLFW--LRWD--CSKS 465

QY 665 KATLGLGGVAVILGAVMAAMGFFSYLGRSSILVILQVVPFVLVSAGADNIFIVLEVQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCLGISNFAATQVLPFELAGVGVDVDFELLAHAFASET 525

QY 725 PRPGBPREVHIGRALGRVAPSMMLCSLSEAIFFLIGALTTPMPANRVTALTSLGLAVILDF 784
DB 526 GQNKRIFFEDRTGECLKRTKRTGASVALTSISNVTFAFMAALPIPALRAFSQAQAVVVVNF 585

QY 785 LLQMSAFVALLSLDSKQEARLDDVCC-----VKQBEL-----PPG 822
DB 586 AMVLLIFAILSMIDLRYRDRRLDIFCCFTSPCVSRVIOVEPQAYTDHONTRVSPPPY 645

QY 823 QGEG----- 827
DB 646 SSHSFAHETQITMGSTVQLRTEYDPHTHVYVTTAEPRSEISVQPVTVTQDTLSCQSPST 705

QY 828 -----LLGPFQKAYAPFLHWHITRGVVLLFLALFGVSL 861
DB 706 SSTRDLLSQFSDSSLHCLPEPCTKWTLSFAEKHYAPFLKPKAKVWVIFLFLGLLGVSL 765

QY 862 YSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVVFVTTLGYNFSSEAGMNAICS 921
DB 766 YGTRVRDGLDITVPRETREYDFIAAQKYPFS-----YNM----- 803

QY 922 SAGCNFSFTQIQYAT-----EFPEQSYLAIPASS-----WVDDFIDM----- 961
DB 804 -----YIVTQKADYPIQHLLYDLHRSFSNVKVMLEENKQLPKMHLHYFRDMLQGLQD 857

QY 962 -----TPSS-----CCRLYISGNKQKFCPSTVNSLNCNCKNSITMGS 1000
DB 858 AFDSDWETGKIMPNYKNGSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913

QY 1001 VRPSVEQPHKYL--PWFLND-----RPNKCPKGGLAAY--STSVNLTSQDQ 1043
DB 914 INPSA--FYILTAWVSNDFVAYAAQANIRPHPEWHDK--ADYMPETRLRIPAAEP 968

QY 1044 VLASRFMAYHFKPLKNSODYTEALRAARELANITA-DLRKVPGTDPAPEVPPYTTNVFY 1102
DB 969 IEYAFQPFYLNGLRDTSDVEAIEKVRTICSNYTSGLSSYPNG-----YPP-----LFW 1018

QY 1103 EGYLTILPEGLPMLSCLVPTFANVSCLLGLDLRGLNLLSIVMLVDTVGFMALWDIS 1162
DB 1019 EGYIGLRHWWLLFISVWLACTFLVCVFLNPNWTAGII--VWVIALMTVELFGMGLIGIK 1077

QY 1163 YNAVSLINLVSAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAYFAGVAMTN 1220
DB 1078 LSAPVPTILASVGIQVEFTHVALAFLTAIGDKN--RRVLALEHMPAPVLDG-AVST 1133

QY 1221 LFGILVLGAKAQIQLIOPFFRLNLLITLLGLLHGLVFLPVLSYVG--PDVNPALALEQ- 1277
DB 1134 LGLVLMAGSEDFIVRYFAVLAAILITLGLVNLGLVLLPVLLSPFGPYEVPSPANGLNRL 1193

QY 1278 --KRAEAAVAVMVASCPNHPSPVSTADNIYVNHSPSGSIKG 1317
DB 1194 PTPSPPPPSVVRFAAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

Search completed: April 7, 2006, 19:23:05
Job time : 63 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:32:55 ; Search time 188 Seconds
(without alignments)
2960.365 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNLPNNGRQF 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	US-10-621-758A-4	Sequence 4, Appli
2	6909	100.0	1332	US-10-663-208A-4	Sequence 4, Appli
3	6909	100.0	1332	US-10-646-301A-4	Sequence 4, Appli
4	6909	100.0	1332	US-10-736-769-4	Sequence 4, Appli
5	6896	99.8	1332	US-10-239-316-9	Sequence 9, Appli
6	6872.5	99.5	1359	US-10-621-758A-4	Sequence 4, Appli
7	6872.5	99.5	1359	US-10-663-208A-4	Sequence 4, Appli
8	6872.5	99.5	1359	US-10-646-301A-4	Sequence 4, Appli
9	6872.5	99.5	1359	US-10-736-769-4	Sequence 4, Appli
10	6536	94.6	1344	US-10-450-763-53052	Sequence 53052, A
11	5421.5	78.5	1331	US-10-621-758A-2	Sequence 2, Appli
12	5421.5	78.5	1331	US-10-663-208A-2	Sequence 2, Appli
13	5421.5	78.5	1331	US-10-646-301A-2	Sequence 2, Appli
14	5421.5	78.5	1331	US-10-736-769-2	Sequence 2, Appli
15	5407	78.3	1333	US-10-621-758A-12	Sequence 12, Appli
16	5407	78.3	1333	US-10-663-208A-12	Sequence 12, Appli
17	5407	78.3	1333	US-10-646-301A-12	Sequence 12, Appli
18	5407	78.3	1333	US-10-736-769-12	Sequence 12, Appli
19	4466	64.6	982	US-10-450-763-53050	Sequence 53050, A
20	2402.5	34.8	1278	US-10-208-731-2	Sequence 2, Appli
21	2402.5	34.8	1278	US-10-741-600-530	Sequence 530, App
22	2402.5	34.8	1278	US-10-741-600-1542	Sequence 1542, App
23	2402.5	34.8	1278	US-10-756-149-4924	Sequence 4924, App
24	2385	34.5	1319	US-10-208-731-4	Sequence 4, Appli
25	2262	32.7	1287	US-11-097-143-12003	Sequence 12003, A
26	1864.5	27.0	1223	US-11-097-143-2679	Sequence 2679, App
27	1329	19.2	1170	US-10-208-731-6	Sequence 6, Appli

28	1078	15.6	1296	4	US-10-208-731-9	Sequence 9, Appli
29	942	13.6	194	4	US-10-239-316-8	Sequence 8, Appli
30	900.5	13.0	492	4	US-10-424-599-211862	Sequence 211862,
31	859.5	12.4	541	4	US-10-424-599-189288	Sequence 189288,
32	651.5	9.4	1447	2	US-08-954-701A-19	Sequence 19, Appli
33	651.5	9.4	1447	3	US-09-898-533-5	Sequence 5, Appli
34	651.5	9.4	1447	3	US-09-754-032-19	Sequence 19, Appli
35	651.5	9.4	1447	4	US-10-421-446-19	Sequence 19, Appli
36	651.5	9.4	1447	4	US-10-791-844-6	Sequence 6, Appli
37	646.5	9.4	1434	2	US-08-954-701A-10	Sequence 10, Appli
38	646.5	9.4	1434	3	US-09-754-032-10	Sequence 10, Appli
39	646.5	9.4	1434	4	US-10-421-446-10	Sequence 10, Appli
40	628	9.1	1296	4	US-10-302-279-60	Sequence 60, Appli
41	598.5	8.7	933	4	US-10-415-934-3	Sequence 3, Appli
42	597	8.6	821	5	US-10-890-776A-4805	Sequence 4805, Ap
43	588	8.5	954	4	US-10-060-756A-3	Sequence 3, Appli
44	588	8.5	954	5	US-10-890-776A-3	Sequence 3, Appli
45	584	8.5	1203	3	US-09-990-046-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-621-758A-4
; Sequence 4, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4

Query Match	100.0%	Score	6909;	DB	4;	Length	1332;
Best Local Similarity	100.0%;	Pred.	No. 0;				
Matches	1332;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MAEAGLRGWLWALLRLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN	60				
Db	1	MAEAGLRGWLWALLRLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN	60				
QY	61	TPARKITGDHLTLLOKICPRLYTGPNTOQCCSAKQVLSLEASLSITKALLTRCPACSDNF	120				
Db	61	TPARKITGDHLTLLOKICPRLYTGPNTOQCCSAKQVLSLEASLSITKALLTRCPACSDNF	120				
QY	121	VNLHCHTCSNPQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSVDSCSRVRPAA	180				
Db	121	VNLHCHTCSNPQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSVDSCSRVRPAA	180				
QY	181	ATLVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDTFHLLRPGQAVGSGIQPLNEGV	240				
Db	181	ATLVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDTFHLLRPGQAVGSGIQPLNEGV	240				
QY	241	ARCNESQGDVATCSQDCCAASCPAIPQALDSTFYLQMPGSLVLIILCSVFAVVTI	300				
Db	241	ARCNESQGDVATCSQDCCAASCPAIPQALDSTFYLQMPGSLVLIILCSVFAVVTI	300				
QY	301	LLVGRVAPARDKSNQVDPKKGTSLSDKLSFSTHTLLQOFFQCGWGTWVASWPLTILVLSV	360				
Db	301	LLVGRVAPARDKSNQVDPKKGTSLSDKLSFSTHTLLQOFFQCGWGTWVASWPLTILVLSV	360				


```
Db 661 MVDKATLGLGVA VVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSLVSGADNIFIVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFLGLALTMPAVRTFALTSGLA 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFLGLALTMPAVRTFALTSGLA 780
Qy 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQBELPPQGGGLLGFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQBELPPQGGGLLGFQKAYAPPL 840
Qy 841 LHWITRGVVLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPFQSYLAIPASSWDDFD 960
Db 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPFQSYLAIPASSWDDFD 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSNLTSQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSNLTSQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFVFPVYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSG 1140
Db 1081 RKVPGTDPAFVFPVYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSG 1140
Qy 1141 NLLSIVMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKTWLER 1200
Db 1141 NLLSIVMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKTWLER 1200
Qy 1201 AKEATISGSAFVAGVAMTNLPGILVLGLAKAQLIQIPFRLNLTLLGLLHGLVFLPV 1260
Db 1201 AKEATISGSAFVAGVAMTNLPGILVLGLAKAQLIQIPFRLNLTLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPRSVSTADNIYVNHSPSGSKGAGA 1320
Db 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPRSVSTADNIYVNHSPSGSKGAGA 1320
Qy 1321 ISNPLPNNGRQF 1332
Db 1321 ISNPLPNNGRQF 1332
```

RESULT 3

```
US-10-646-301A-4
; Sequence 4, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-4
```

```
Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MABAGLWLLWALLRLAQSEPTTIIHQPGYCAFYDECGKNPELSGLMTLSNVSCLN 60
Db 1 MABAGLWLLWALLRLAQSEPTTIIHQPGYCAFYDECGKNPELSGLMTLSNVSCLN 60
Qy 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKOLVSLSEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKOLVSLSEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNPQSLFPIVTRVAQGLGAGQLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHTCSNPQSLFPIVTRVAQGLGAGQLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAVGTWCGVYGSALCNAQRLNFGQDGTGNGLAPLDITFHLLERQQAQVSGIQLNEGV 240
Db 181 ATLAVGTWCGVYGSALCNAQRLNFGQDGTGNGLAPLDITFHLLERQQAQVSGIQLNEGV 240
Qy 241 ARCNEQGDVATCSQDCAASCAPAIARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCAPAIARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
Qy 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSPTSTHTLLQOPFGCWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSPTSTHTLLQOPFGCWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQASEKAFHQHFGPPFRTRNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQASEKAFHQHFGPPFRTRNQVILTAPNRSY 420
Qy 421 RYDSLGLLGPKNFGSILDLLELEQLERLHLQWSPAEQARNISLODICVAPLNPDNT 480
Db 421 RYDSLGLLGPKNFGSILDLLELEQLERLHLQWSPAEQARNISLODICVAPLNPDNT 480
Qy 481 SLYDCINSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTPKDGFTALAL 540
Db 481 SLYDCINSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTPKDGFTALAL 540
Qy 541 SCMDYGPVPPFLAIGYKGYKDYSEALIMTFSLNYPAGDPRLAQKLEAFLEEM 600
Db 541 SCMDYGPVPPFLAIGYKGYKDYSEALIMTFSLNYPAGDPRLAQKLEAFLEEM 600
Qy 601 RAFORMAGMFOVTFATERSLEDEINRTTAEPLPATSIIYVILYISLALSGYSWSRV 660
Db 601 RAFORMAGMFOVTFATERSLEDEINRTTAEPLPATSIIYVILYISLALSGYSWSRV 660
Qy 661 MVDKATLGLGVA VVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSLVSGADNIFIVLE 720
Db 661 MVDKATLGLGVA VVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSLVSGADNIFIVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFLGLALTMPAVRTFALTSGLA 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFLGLALTMPAVRTFALTSGLA 780
Qy 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQBELPPQGGGLLGFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQBELPPQGGGLLGFQKAYAPPL 840
Qy 841 LHWITRGVVLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPFQSYLAIPASSWDDFD 960
Db 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPFQSYLAIPASSWDDFD 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSNLTSQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
```

Db	1021	NIKCPKGLAAYSTSVNLTSQVLAGRFMAHYKPLKNSQDYTEALRAARELANITADL	1080
Qy	1081	RKVPGTDPAPFVFPYTTINNVFYOVLTLPGELFMLSCLVPTFAVSCLLIIGLDRSGLL	1140
Db	1081	RKVPGTDPAPFVFPYTTINNVFYOVLTLPGELFMLSCLVPTFAVSCLLIIGLDRSGLL	1140
Qy	1141	NLLSIVMLVDTVGFPMALWDISYNVSLINLVSAVGMSEVFSHITRSFAISTKPTWLER	1200
Db	1141	NLLSIVMLVDTVGFPMALWDISYNVSLINLVSAVGMSEVFSHITRSFAISTKPTWLER	1200
Qy	1201	AKEATISGSAVFAGVAMTNLPGILVGLAKAQLIQIIEFFRLNLLITLGLLHGLVFLPV	1260
Db	1201	AKEATISGSAVFAGVAMTNLPGILVGLAKAQLIQIIEFFRLNLLITLGLLHGLVFLPV	1260
Qy	1261	ILSYVGPDPVPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA	1320
Db	1261	ILSYVGPDPVPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA	1320
Qy	1321	ISNFLPNNGRQF 1332	
Db	1321	ISNFLPNNGRQF 1332	
RESULT 4			
US-10-736-769-4			
; Sequence 4, Application US/10736769			
; Publication No. US20040161838A1			
; GENERAL INFORMATION:			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Wang, Luquan			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JB01603-K3-US			
; CURRENT APPLICATION NUMBER: US/10/736,769			
; CURRENT FILING DATE: 2003-12-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; PRIOR APPLICATION NUMBER: 10/663,208			
; PRIOR FILING DATE: 2003-09-16			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 4			
; LENGTH: 1332			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-736-769-4			
Query Match 100.0%; Score 6909; DB 4; Length 1332;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECCKNPGLSGSLMTLSNVSCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECCKNPGLSGSLMTLSNVSCLSN	60
Qy	61	TPARKITGDHLILLOKICPLRYTGPNTQACCSAKQLVLEASISITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLILLOKICPLRYTGPNTQACCSAKQLVLEASISITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSFNQSLFNVTRVAQLGAGQLPAVVAEYAFQHSFAEQSDSCSRVRVPA	180
Db	121	VNLHCHNTCSFNQSLFNVTRVAQLGAGQLPAVVAEYAFQHSFAEQSDSCSRVRVPA	180
Qy	181	ATLAVGTWCGVYGSAALNAQRLNFGDGTGNGLAPLDTITPHLEPGQVSGIQPLNEGV	240
Db	181	ATLAVGTWCGVYGSAALNAQRLNFGDGTGNGLAPLDTITPHLEPGQVSGIQPLNEGV	240

```
Db      1321 ISNPLPNNGRQF 1332
RESULT 5
US-10-239-316-9
; Sequence 9, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same And
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-9
Query Match          99.8%; Score 6896; DB 4; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 MAEAGLRGWLWALLRLAQSBPYTHIQGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
Db      1 MAEAGLRGWLWALLRLAQSBPYTHIQGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
QY      61 TPARKITGDHLILLQKICPRLYTGNTQACCSAKQLVLSLEASISITKALLTRCPACSNF 120
Db      61 TPARKITGDHLILLQKICPRLYTGNTQACCSAKQLVLSLEASISITKALLTRCPACSNF 120
QY      121 VNLHCHTNCSPNQLFINVTRVAQAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
Db      121 VNLHCHTNCSPNQLFINVTRVAQAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
QY      181 ATLVGTMCGVYGSALCNAQRLNFGQDTGNGLAFLDITFHLLEPCQAVGSGIQPLNEGV 240
Db      181 ATLVGTMCGVYGSALCNAQRLNFGQDTGNGLAFLDITFHLLEPCQAVGSGIQPLNEGV 240
QY      241 ARCNEQGDVATCQDCQCAAPARQALDSTFYLGQMPGSLVLIILICSVFAVVTI 300
Db      241 ARCNEQGDVATCQDCQCAAPARQALDSTFYLGQMPGSLVLIILICSVFAVVTI 300
QY      301 LLVGRFVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360
Db      301 LLVGRFVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360
QY      361 IPVVALAAGLVTELTDDPVELWSAPNSQARSEKAFDQHFQFFRTNQVILTAPNRSY 420
Db      361 IPVVALAAGLVTELTDDPVELWSAPNSQARSEKAFDQHFQFFRTNQVILTAPNRSY 420
QY      421 RYDLSLLGPNKPSGILDLDLLELELQERLRLQVSPQAQRNLSLQDICYAPLNPDNT 480
Db      421 RYDLSLLGPNKPSGILDLDLLELELQERLRLQVSPQAQRNLSLQDICYAPLNPDNT 480
QY      481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPYLCANAPLTPEKGTALAL 540
Db      481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPYLCANAPLTPEKGTALAL 540
QY      541 SCWADYGAVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWEAFLEEM 600
Db      541 SCWADYGAVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWEAFLEEM 600
QY      601 RAFORMMAGMFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYGSWSRV 660
Db      601 RAFORMMAGMFQVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYGSWSRV 660
661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVL 720
661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVL 720
721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFPLGALTMPAVRTTALTSGLA 780
721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFPLGALTMPAVRTTALTSGLA 780
781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQELPPPGQGGSLILGPFQKAYAPFL 840
781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQELPPPGQGGSLILGPFQKAYAPFL 840
841 LHMWTRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDVFLFLNRYFEVGAPV 900
841 LHMWTRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDVFLFLNRYFEVGAPV 900
901 YFVTTLYGNFSEAGMNAICSSAGCNFSTFKIQYATEFPFQSYLAIPASSWVDDFIDW 960
901 YFVTTLYGNFSEAGMNAICSSAGCNFSTFKIQYATEFPFQSYLAIPASSWVDDFIDW 960
961 LTPSSCCRLYISGPNKDKPCSTVNSLNCIKNCSITMGSVRPSVEQHKYLPWFINDRP 1020
961 LTPSSCCRLYISGPNKDKPCSTVNSLNCIKNCSITMGSVRPSVEQHKYLPWFINDRP 1020
1021 NIKCPKGLAAYSTSVNLTSDQVLAASFMAVHKPLKNSQDYTEALRAARELANITADL 1080
1021 NIKCPKGLAAYSTSVNLTSDQVLAASFMAVHKPLKNSQDYTEALRAARELANITADL 1080
1081 RKVPGTDPAFEPVPTIITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGL 1140
1081 RKVPGTDPAFEPVPTIITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGL 1140
1141 NLLSIVMLVDTVGFMAWLDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200
1141 NLLSIVMLVDTVGFMAWLDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200
1201 AKEATISMGSAVPAVAMTNLPGILVLAKAQLIQIPFFRLNLITLILGLHGLVFLPV 1260
1201 AKEATISMGSAVPAVAMTNLPGILVLAKAQLIQIPFFRLNLITLILGLHGLVFLPV 1260
1261 ILSYVGPDPVPALALEQKRAEEAAMVAVASCPNHPSPRSTADNIVVNHSPFGSINGAGA 1320
1261 ILSYVGPDPVPALALEQKRAEEAAMVAVASCPNHPSPRSTADNIVVNHSPFGSINGAGA 1320
1321 ISNPLPNNGRQF 1332
1321 ISNPLPNNGRQF 1332
RESULT 6
US-10-621-758A-44
; Sequence 44, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Luquan
; APPLICANT: Altman, Scott W
; APPLICANT: Graziano, Michael
; TITLE OF INVENTION: NPC1L1
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-44
```


Query Match		99.5%;	Score 6872.5;	DB 4;	Length 1359;	
Best Local Similarity		97.9%;	Pred. No. 0;			
Matches 1330;		Conservative	0;	Mismatches	2;	Indels 27; Gaps 1;
Qy	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECGKNPELGSGLMTLSNVCSLSN	60			
Db	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECGKNPELGSGLMTLSNVCSLSN	60			
Qy	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Db	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180			
Db	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180			
Qy	181	ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240			
Db	181	ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240			
Qy	241	ARCNESQGDVATCSQDCAACSPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300			
Db	241	ARCNESQGDVATCSQDCAACSPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300			
Qy	301	LLVGFVAPARDKSMVDPKGTSLSDKLSFSHTLLGQFFQGWGTWASWPLTILVLSV	360			
Db	301	LLVGFVAPARDKSMVDPKGTSLSDKLSFSHTLLGQFFQGWGTWASWPLTILVLSV	360			
Qy	361	IPVALLAAGLVFTELTDVPELWASNSQARSEKAFHDOHFGPFPTNQTIVITAPNRSY	420			
Db	361	IPVALLAAGLVFTELTDVPELWASNSQARSEKAFHDOHFGPFPTNQTIVITAPNRSY	420			
Qy	421	RYDSLLGPKNFSGIILLOLLELLELQERLRLHQLVWSPQAQRNLSLODICVAPLNPDNT	480			
Db	421	RYDSLLGPKNFSGIILLOLLELLELQERLRLHQLVWSPQAQRNLSLODICVAPLNPDNT	480			
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTKOGTALAL	540			
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTKOGTALAL	540			
Qy	541	SCWADYGAPVFPFLAIGYKGDYSAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLEM	600			
Db	541	SCWADYGAPVFPFLAIGYKGDYSAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLEM	600			
Qy	601	RAFORMMAGMFQVTFMAERSLEDEINRTTAEDLPFATSIVITFLYISLALGYSYSSWSRV	660			
Db	601	RAFORMMAGMFQVTFMAERSLEDEINRTTAEDLPFATSIVITFLYISLALGYSYSSWSRV	660			
Qy	661	MVDSKATLGLGGVAVVLGAVMAANGFYSYLGIRSSVLQVLPFLVLSVGDADNIFIFVLE	720			
Db	661	MVDSKATLGLGGVAVVLGAVMAANGFYSYLGIRSSVLQVLPFLVLSVGDADNIFIFVLE	720			
Qy	721	YQRLPRPGRPREVHIGRALGRVAPSMMLCSLSEACFFILGALTPMPAVRTFALTSLAV	780			
Db	721	YQRLPRPGRPREVHIGRALGRVAPSMMLCSLSEACFFILGALTPMPAVRTFALTSLAV	780			
Qy	781	ILDFLLQMSAFVALLSLDSKREASRLDVCCKVPQELPPPGQEGLLILFFQKAYAPFL	840			
Db	781	ILDFLLQMSAFVALLSLDSKREASRLDVCCKVPQELPPPGQEGLLILFFQKAYAPFL	840			
Qy	841	LHWITRGVLLFLALFGVSLYSNCHISVGLDQSLALPKDSYLDLYFLNRYFEVGA	900			
Db	841	LHWITRGVLLFLALFGVSLYSNCHISVGLDQSLALPKDSYLDLYFLNRYFEVGA	900			
Qy	901	YFVTTLGNFSSAGMNAICSSAGCNFSTFKIQYATEFPQSYLAIPASSWDDFDID	960			
Db	901	YFVTTLGNFSSAGMNAICSSAGCNFSTFKIQYATEFPQSYLAIPASSWDDFDID	960			
Qy	961	LTPSSCCRLYISGPNKDFCSPTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP	1020			
Db	961	LTPSSCCRLYISGPNKDFCSPTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP	1020			
Qy	1021	NIKCPKGLAAYSTSVNLJTSQGVL-----ASRFMAYH	1053			
Db	1021	NIKCPKGLAAYSTSVNLJTSQGVL-----ASRFMAYH	1080			
Qy	1054	KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNNFYEQYLTILPEGL	1113			
Db	1081	KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNNFYEQYLTILPEGL	1140			
Qy	1114	FMLSCLVPTFAVSCLLGLDLSGLNLALSIVMLVDTVGFMAWLDISYNAVSLINLVS	1173			
Db	1141	FMLSCLVPTFAVSCLLGLDLSGLNLALSIVMLVDTVGFMAWLDISYNAVSLINLVS	1200			
Qy	1174	AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1233			
Db	1201	AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1260			
Qy	1234	LIQIFPFRMLNLLITLLGLLHGLVFLPVILSVGPDVNPALALEQKRAEEAAVAAVMVASCP	1293			
Db	1261	LIQIFPFRMLNLLITLLGLLHGLVFLPVILSVGPDVNPALALEQKRAEEAAVAAVMVASCP	1320			
Qy	1294	NHPSRVSTADNIYVNHSPFSGIKGAGAISNPLPNNGRQF	1332			
Db	1321	NHPSRVSTADNIYVNHSPFSGIKGAGAISNPLPNNGRQF	1359			
RESULT 7						
US-10-663-208A-44						
; Sequence 44, Application US/10663208A						
; Publication No. US20040132058A1						
; GENERAL INFORMATION:						
; APPLICANT: Wang, Luquan						
; APPLICANT: Graziano, Michael						
; APPLICANT: Murgolo, Nick						
; TITLE OF INVENTION: NPC11A (NPC3) AND METHODS OF USE THEREOF						
; FILE REFERENCE: JB01603K2 US						
; CURRENT APPLICATION NUMBER: US/10/663,208A						
; PRIOR FILING DATE: 2003-09-16						
; PRIOR FILING DATE: 2002-07-19						
; PRIOR APPLICATION NUMBER: 10/621,758						
; PRIOR FILING DATE: 2003-07-17						
; PRIOR APPLICATION NUMBER: 10/646,301						
; PRIOR FILING DATE: 2003-08-22						
; NUMBER OF SEQ ID NOS: 50						
; SOFTWARE: PatentIn version 3.1						
; SEQ ID NO 44						
; LENGTH: 1359						
; TYPE: PRT						
; ORGANISM: Homo sapiens						
US-10-663-208A-44						
Query Match						
Best Local Similarity						
Matches 1330;						
Conservative						
0;						
Mismatches						
2;						
Indels						
27;						
Gaps						
1;						
Qy	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECGKNPELGSGLMTLSNVCSLSN	60			
Db	1	MAEAGLRGWLWALLRLAQAQSEPYTHIQPGYCAFYDECGKNPELGSGLMTLSNVCSLSN	60			
Qy	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Db	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180			
Db	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180			
Qy	181	ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240			
Db	181	ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240			
Qy	241	ARCNESQGDVATCSQDCAACSPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300			
Db	241	ARCNESQGDVATCSQDCAACSPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300			
Qy	301	LLVGFVAPARDKSMVDPKGTSLSDKLSFSHTLLGQFFQGWGTWASWPLTILVLSV	360			
Db	301	LLVGFVAPARDKSMVDPKGTSLSDKLSFSHTLLGQFFQGWGTWASWPLTILVLSV	360			
Qy	361	IPVALLAAGLVFTELTDVPELWASNSQARSEKAFHDOHFGPFPTNQTIVITAPNRSY	420			
Db	361	IPVALLAAGLVFTELTDVPELWASNSQARSEKAFHDOHFGPFPTNQTIVITAPNRSY	420			
Qy	421	RYDSLLGPKNFSGIILLOLLELLELQERLRLHQLVWSPQAQRNLSLODICVAPLNPDNT	480			
Db	421	RYDSLLGPKNFSGIILLOLLELLELQERLRLHQLVWSPQAQRNLSLODICVAPLNPDNT	480			
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTKOGTALAL	540			
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTKOGTALAL	540			
Qy	541	SCWADYGAPVFPFLAIGYKGDYSAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLEM	600			
Db	541	SCWADYGAPVFPFLAIGYKGDYSAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLEM	600			
Qy	601	RAFORMMAGMFQVTFMAERSLEDEINRTTAEDLPFATSIVITFLYISLALGYSYSSWSRV	660			
Db	601	RAFORMMAGMFQVTFMAERSLEDEINRTTAEDLPFATSIVITFLYISLALGYSYSSWSRV	660			
Qy	661	MVDSKATLGLGGVAVVLGAVMAANGFYSYLGIRSSVLQVLPFLVLSVGDADNIFIFVLE	720			
Db	661	MVDSKATLGLGGVAVVLGAVMAANGFYSYLGIRSSVLQVLPFLVLSVGDADNIFIFVLE	720			
Qy	721	YQRLPRPGRPREVHIGRALGRVAPSMMLCSLSEACFFILGALTPMPAVRTFALTSLAV	780			
Db	721	YQRLPRPGRPREVHIGRALGRVAPSMMLCSLSEACFFILGALTPMPAVRTFALTSLAV	780			
Qy	781	ILDFLLQMSAFVALLSLDSKREASRLDVCCKVPQELPPPGQEGLLILFFQKAYAPFL	840			
Db	781	ILDFLLQMSAFVALLSLDSKREASRLDVCCKVPQELPPPGQEGLLILFFQKAYAPFL	840			
Qy	841	LHWITRGVLLFLALFGVSLYSNCHISVGLDQSLALPKDSYLDLYFLNRYFEVGA	900			
Db	841	LHWITRGVLLFLALFGVSLYSNCHISVGLDQSLALPKDSYLDLYFLNRYFEVGA	900			
Qy	901	YFVTTLGNFSSAGMNAICSSAGCNFSTFKIQYATEFPQSYLAIPASSWDDFDID	960			
Db	901	YFVTTLGNFSSAGMNAICSSAGCNFSTFKIQYATEFPQSYLAIPASSWDDFDID	960			
Qy	961	LTPSSCCRLYISGPNKDFCSPTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP	1020			
Db	961	LTPSSCCRLYISGPNKDFCSPTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP	1020			

Db 241 ARCNESQDDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGFFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGFFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLDDLELLELQERLHLQVSPQAQRNLSQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDDLELLELQERLHLQVSPQAQRNLSQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
QY 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSRV 660
Db 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAWGPFSYLGIRSSLVILQVVPFLVSVGADNIPFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAWGPFSYLGIRSSLVILQVVPFLVSVGADNIPFVLE 720
QY 721 YQRLPRPCEPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRPCEPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLSKROEASRLDVCCCKPOELPPQCGEGILLGPFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLSKROEASRLDVCCCKPOELPPQCGEGILLGPFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAVP 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAVP 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQVATEFPBQSYLAIPASSWVDVDFIDW 960
Db 901 YFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQVATEFPBQSYLAIPASSWVDVDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSITMGSVRPSVEQPHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSITMGSVRPSVEQPHKYLFWFLNDRP 1020
QY 1021 NIKCPKGLAAVSTSVNLTSDGOVL-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAVSTSVNLTSDGOVLDTVAILSPRLEYSTISAHNCNLYLLDSASRPMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANIATDLRKVPGTDPAFEPFPTITNVFYEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANIATDLRKVPGTDPAFEPFPTITNVFYEQVLTILPEGL 1140
QY 1114 FMLSICLVPTFAVSCLLGLDLSRGLNLLSIVMILVDTVGFMAWWDISYNAVSLINLVS 1173
Db 1141 FMLSICLVPTFAVSCLLGLDLSRGLNLLSIVMILVDTVGFMAWWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHLTRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1233
Db 1201 AVGMSVEFVSHLTRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIFPFRNLNLTILGLLHGLVFLPVLISYVGPVDPNPPALAEQKRAEAAVAVWVASC 1293
Db 1261 LIQIFPFRNLNLTILGLLHGLVFLPVLISYVGPVDPNPPALAEQKRAEAAVAVWVASC 1320
QY 1294 NHPSRVSTADNTIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1332
|||||

Db 1321 NHPSRVSTADNTIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1359
RESULT 8
US-10-646-301A-44
; Sequence 44, Application US/10646301A
; Publication No. US2004040137467A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Luquan
; APPLICANT: Altmann, Scott W
; APPLICANT: Graziano, Michael
; APPLICANT: Murgo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-KI-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-646-301A-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLN 60
QY 61 TPARKITGDHLLILQKICPRLYTGPTNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGPTNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSNQSIFINVTVAQILGAGQLPAVAYEAPYQHSFAPQSVDSCSRVRPAA 180
Db 121 VNLHCHTCSNQSIFINVTVAQILGAGQLPAVAYEAPYQHSFAPQSVDSCSRVRPAA 180
QY 181 ATLAVGTWCVGVGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQQA VSGGIQIPLNEGV 240
Db 181 ATLAVGTWCVGVGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQQA VSGGIQIPLNEGV 240
QY 241 ARCNESQDDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNESQDDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGFFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGFFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLDDLELLELQERLHLQVSPQAQRNLSQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDDLELLELQERLHLQVSPQAQRNLSQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
QY 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSRV 660

Db 601 RAPORMAGMFOVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGSSSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720
Qy 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSLGLAV 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPPGQGBGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPPGQGBGLLGFQKAYAPFL 840
Qy 841 LHMWTRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHMWTRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSBAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
Db 901 YFVTTILGYNFSBAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFLNDRP 1020
Qy 1021 NIKCPKGGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
Db 1021 NIKCPKGGLAAYSTSVNLTSDQVLTVAILSPRLEYSCTTSAHCLNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSODYTEALRAARELAANITADLRKVPCTDPAFVFPVYTI TNVFEQVLTILPGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPVYTI TNVFEQVLTILPGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVLIMLVDTVGFMAWLDISYNAVSLNLYS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVLIMLVDTVGFMAWLGISYNAVSLNLYS 1200
Qy 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVGLAKAQ 1260
Qy 1234 LIQIFPRLNLLITLGLHLGLVFLPVILSYVGPVNPALAEOKRAEAAVAVWVASC 1293
Db 1261 LIQIFPRLNLLITLGLHLGLVFLPVILSYVGPVNPALAEOKRAEAAVAVWVASC 1320
Qy 1294 NHPSRVSTADNIYVNHSPFGSIKAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAI SNFLPNNGRQF 1359

RESULT 9
US-10-736-769-44
; Sequence 44, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10736, 769
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16

; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLIIWALLRLAQSEPYTHIQPGYCAFYDECCKNPGLSGSLMTLSNVCLSN 60
Db 1 MAEAGLRGWLIIWALLRLAQSEPYTHIQPGYCAFYDECCKNPGLSGSLMTLSNVCLSN 60
Qy 61 TPARKITGDHIILOKI CPHLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSNF 120
Db 61 TPARKITGDHIILOKI CPHLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAVGTCGYGSGALCNAQRWLFQDGTGNGLAPLDITFHLLEPGQAVGSGIOLNEGV 240
Db 181 ATLAVGTCGYGSGALCNAQRWLFQDGTGNGLAPLDITFHLLEPGQAVGSGIOLNEGV 240
Qy 241 ARCHESQDDVATCSQDCAASCALARPQALDSTFYLQMPGSLVLIILICSVPAVVTI 300
Db 241 ARCHESQDDVATCSQDCAASCALARPQALDSTFYLQMPGSLVLIILICSVPAVVTI 300
Qy 301 LLVGFVAPARDKSMVDPKKGTSLSDKLSFTHTLLGQFFQGMGTWVASMPLTILVLSV 360
Db 301 LLVGFVAPARDKSMVDPKKGTSLSDKLSFTHTLLGQFFQGMGTWVASMPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGPPFRTHQVILTAENRSY 420
Db 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGPPFRTHQVILTAENRSY 420
Qy 421 RYDLSLLGPKRPSGTLDDLLLELLEQLERHLQVWSPQARNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKRPSGTLDDLLLELLEQLERHLQVWSPQARNISLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTAL 540
Qy 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEM 600
Qy 601 RAFORMMAGMFOVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGSSSRV 660
Db 601 RAFORMMAGMFOVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGSSSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGADNIFIFVLE 720
Qy 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSLGLAV 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPPGQGBGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPPGQGBGLLGFQKAYAPFL 840
Qy 841 LHMWTRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHMWTRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSBAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960

```

901 YFVTITGYNFSSAGMNAICSSAGCNFFSTQKIYATFPPQSYLAIPASSWVDDFIDW 960
961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVQFHKYLPWFNDRP 1020
961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVQFHKYLPWFNDRP 1020
1021 NIKCPKGGAAVSTSVNLTSDDQVL-----ASRFMAYH 1053
1021 NIKCPKGGAAVSTSVNLTSDDQVLDTVAILSPRLSEYSGTISAHCNLYLLDSASRFMAYH 1080
1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFEPYPTITNVFVEQYLTILPEGL 1113
1081 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFEPYPTITNVFVEQYLTILPEGL 1140
1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDTVGFMAWMDISYNAVSLINLVS 1173
1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDTVGFMAWLGISYNAVSLINLVS 1200
1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAFAGVAMTNLPGILVLGLAKAQ 1233
1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAFAGVAMTNLPGILVLGLAKAQ 1260
1234 LIQIFPFRNLITLGLHGLVFLPVILSYGPDVNPALALEQKRAEBAVAAMVASCP 1293
1261 LIQIFPFRNLITLGLHGLVFLPVILSYGPDVNPALALEQKRAEBAVAAMVASCP 1320
1294 NHPSRVSTADNIVNHSFEGSIKGAGAINFPLNNGRQF 1332
1321 NHPSRVSTADNIVNHSFEGSIKGAGAINFPLNNGRQF 1359

RESULT 10
US-10-450-763-53052
; Sequence 53052, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53052
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (901)..(917)
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,
; accession number PR00543D, p-value=1.355e-09, raw score of 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1344)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53052

Query Match 94.6%; Score 6536; DB 5; Length 1344;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

QY 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKPNELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKPNELSGSLMTLSNVSCLSN 60
```

```

61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120
61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120
121 VNLHCHTCSNQSILFINTRVAQLGAGQLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
121 VNLHCHTCSNQSILFINTRVAQLGAGQLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
181 ATLAVGTCGVVGSALCNAQRWLNFGDGTNGNLAPLDITFHLLEPQAVGSGIQLNEGV 240
181 ATLAVGTCGVVGSALCNAQRWLNFGDGTNGNLAPLDITFHLLEPQAVGSGIQLNEGV 240
241 ARCNSQGDVATCSCQDCAASCP--AIARPAQLDSTFYLGQMPGSLVLIILCSVPFVV 298
241 ARCNSQGDVATCSCQDCAASCPXAPRPPSTW--RQMPGSLVLIILCSVPFVV 298
299 TILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTILGQFQGHGTWVASHPLILVL 358
299 TILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTILGQFQGHGTWVASHPLILVL 358
359 SVIPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQPPFRTNQVILTAPNRS 418
359 SVIPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQPPFRTNQVILTAPNRS 418
419 SYRYSLLGPKNPFSGIILDLLLELLEQLERLHLQVWSPQAQRNISLQDICYAPLNDP 478
419 SYRYSLLGPKNPFSGIILDLLLELLEQLERLHLQVWSPQAQRNISLQDICYAPLNDP 478
479 NTSLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFYCANAPITFKDGTAL 538
479 NTSLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFYCANAPITFKDGTAL 538
539 ALSOMADYGAPVPPFLAIGYGGKDYSEAEALIMTFLSNYPAGDRLQAQKLWEAFLE 598
539 ALSOMADYGAPVPPFLAIGYGGKDYSEAEALIMTFLSNYPAGDRLQAQKLWEAFLE 598
599 EMRAFQRRMAGMFQVTFMAERSLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSSWS 658
599 EMRAFQRRMAGMFQVTFMAERSLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSSWS 658
659 RVWDSKATLGLGGVAVVLGAVMAAMGPPSYGIRSSVLILQVVPPLVLSVAGDNIFIFV 718
659 RVWDSKATLGLGGVAVVLGAVMAAMGPPSYGIRSSVLILQVVPPLVLSVAGDNIFIFV 718
719 LEYQELPRRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTTPMPAVRTALTSG 778
719 LEYQELPRRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTTPMPAVRTALTSG 778
779 AVILDFLLQMSAFVALLSLDSKQBSRLDVCCVKPQELPPPGQEGILLGFFQKAYAP 838
779 AVILDFLLQMSAFVALLSLDSKQBSRLDVCCVKPQELPPPGQEGILLGFFQKAYAP 838
839 FLLHWITRGV-----LFLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYF 894
839 FLLHWITRGV-----LFLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYF 894
895 EVGAPVYFVTTLYGYNFSSEAGMNAICSSAGCNFFSTQKIYATFPPQSYLAIPASSWV 954
895 EVGAPVYFVTTLYGYNFSSEAGMNAICSSAGCNFFSTQKIYATFPPQSYLAIPASSWV 958
955 DDFIDWLTSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVQFHKYLPW 1014
955 DDFIDWLTSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVQFHKYLPW 1018
1015 FLNDRPNIKCPKGGAAVSTSVNLTSDDQVTLASRFMAYHKPLKNSQDYTEALRAARELA 1074
1015 FLNDRPNIKCPKGGAAVSTSVNLTSDDQVTLASRFMAYHKPLKNSQDYTEALRAARELA 1078
1075 NITADLRKVGCTDPAFEPYPTITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLD 1134
1075 NITADLRKVGCTDPAFEPYPTITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLD 1138
1135 LRSGLNLLSIWMLVDTVGFMAWMDISYNAVSLINLVSAGMSVEFVSHITRSPAISTK 1194
```

Db 1139 LRSGLNLLSVMLVDTVGFMAWGLSYNAVSLINLSAVGMSVEFVSHITRSPASTK 1198
Qy 1195 PTWLERAKEATISMGSAFAGVAMTNLPGLILVLGLAKAQLTIQIFFRNLNLLITLLGLLHG 1254
Db 1199 PTWLERAKEATISMGSAFAGVAMTNLPGLILVLGLAKAQLTIQIFFRNLNLLITLLGLLHG 1258
Qy 1255 LVFLPVILSYGPDVNPALALEQKRAEA-----VAAMVWASCPNHPHSRVST 1301
Db 1259 LVFLPVILSYGPDVNPALALEQKRAEGGSGGGLPQAQITTPPESPOLTTMSMT 1313
RESULT 11
US-10-621-758A-2
; Sequence 2, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2
Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPGLSGMLTSLNSVCLSN 60
Db 1 MAEAWL-GWLLWALLLSAAGELYTPKHEAGVCTFYEECGKNPGLSGMLTSLNSVCLSN 59
Qy 61 TPARKITGHLILLOKICRPLYTGPNTO-ACCSAKOLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGHLILLOKICRPLYTGPNNTTACCSTKQLLSLESSMSITKALLTRCPACSDN 119
Qy 120 FVNLHCHNTCSNQSFLINVTVAOLGAGOLPAVAYEAFYQHSFAEQYDSCSRVRPA 179
Db 120 FVNLHCHNTCSNQSFLINVTVAOLGAGOLPAVAYEAFYQHSFAEQYDSCSRVRPA 179
Qy 180 AATLAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDTIFHLLEPGQAVGSGIQPLNKG 239
Db 180 AASLAVGSMCGVYGSALCNAQRWLNFGQDTGNGLAPLDTIFHLLEPGQALPDGIQPLNGK 239
Qy 240 VARNESQGDVATCSQDCAAPARQALDSTFYLGQMPGSLVLIILCSVAVVT 299
Db 240 IAPCNESQGDSDSAVCSQDCAACPVIIPPEALRPSFYMGPMGWLALIIIFAVFVLLS 299
Qy 300 ILVGFVAVARADSKMVDPKGTSLSDKLSFSHTLLGQFQGWGTWVASWPLTILVLS 359
Db 300 AVLRLRVNRRNKNKAEQGQAPKLPKHKLSPHILGRFQNGWTRVWASWPLTVLALS 359
Qy 360 VIPVALAAGLVFTLTDTPVELWSPNSQARSEKAFHDQHPFPFRTNQVILTAENRSS 419
Db 360 FIVVIALAAGLTFTLTDTPVELWSPNSQARSEKSPHDSHFPGFPFRTNQIFVTARNRSS 419
Qy 420 YRYDSLLGPKNFSGIILDLLELLEQLRHLQVWSPQARNISLQDICVAPLNPDN 479
Db 420 YKYSLLGSKNFSGIILDLLELLEQLRHLQVWSPQARNISLQDICVAPLNPDN 479
Qy 480 TSLYDCCINSLSLQYFQNNRTLLMLTANQTLNGQTSQVMDKDFLYCANAPLTFKDGTSLA 539

Db 480 TSLDCCVNSLLQYFQNNRTLLMLTANQTLNGQTSQVMDKDFLYCANAPLTFKDGTSLA 539
Qy 540 LSCWADYGAPVFPFLAIGGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAQKLEAELEE 599
Db 540 LSCWADYGAPVFPFLAIGGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAQKLEAELEE 599
Qy 600 MRAFORMACMFQVTFATERSLDEINRTTAEDLPFATSYIVIFLYISLALSGSYSWSR 659
Db 600 MESFORNTSKQVAFSAERSLEINRTTIQDLFPFVAVSYIIVFLYISLALSGSYSCSR 659
Qy 660 VMVDSKATLGLGGVAVVGLGAVMAAMGFYSGLTRSSLVILQVPPFLVLGVAGDNIPIFVL 719
Db 660 VAVESKATLGLGGVIVVGLGAVLAAMGFYSYLGVPSSLVIIQVWPFLVAVGADNIFPVL 719
Qy 720 EYQELPRRGEPRVHIGRALGRVAPSMALCSLSEALCFFLGALTMPAVRTFALTSGLA 779
Db 720 EYQELPRMPGEQREAHIGRTLGSVAPSMALCSLSEALCFFLGALTMPAVRTFALTSGLA 779
Qy 780 VILDFLLQMSAFVALLSLDSKQEAASLDVCCVKKQELPPPQGGEGLLGFFOKAYAPF 839
Db 780 IILDFLLQMTAFVALLSLDSKQEAASRPDLCCPSTRLPPPKKEGELLRRFFRKIYAPF 839
Qy 840 LLHWITRGVLLLFALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRVFEYCAP 899
Db 840 LLHRFIRPVVMLLFLTLFGANLYLMCNINVGLDOELALPKDSYLLDYFLFLNRVLEVGPP 899
Qy 900 VYFVTTLGYNFSSEAGNATCSSAGCNFSFTQKIQVATEPPEOSYLAIPASSWVDDPID 959
Db 900 VYFVTTSFGNFSSEAGNATCSSAGCKSFSUTQKIQVASEPFDOSYVIAASSWVDDPID 959
Qy 960 WLTP-SSCCRLYISGPNKDFCPSVTNSLNCNCKMSITMGSVRSPSVQFHKYLPWFUND 1018
Db 960 WLTPSSCCRLYIRGPHKDFCPSVTDTSFNCLKNCNRTLGPVPTABQPHKYPWFUND 1019
Qy 1019 RPNKCPKGLAAAYSTSVNLSDQVLAASRMAYHKLKNSODTEALRAARELAANITA 1078
Db 1020 PENIRCPKGLAAAYRTSVNLSSDQVIAQSPMAYHKLKNSQDTEALRASRLAANITA 1079
Qy 1079 DLKRVGTDPAFEPFVPTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSG 1138
Db 1080 DLKRVGTDPAFEPFVPTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSG 1139
Qy 1139 LNLISLIVMLVDTVGFMAWGLSYNAVSLINLSAVGMSVEFVSHITRSPASTKPTL 1198
Db 1140 LNLISLIVMLVDTVGFMAWGLSYNAVSLINLSAVGMSVEFVSHITRSPASTKPTL 1199
Qy 1199 ERAKATISMGSAFAGVAMTNLPGLILVLGLAKAQLTIQIFFRNLNLLITLLGLLHVFL 1258
Db 1200 ERAKATISMGSAFAGVAMTNLPGLILVLGLAKAQLTIQIFFRNLNLLITLLGLLHVFL 1259
Qy 1259 PVILSYGPDVNPALALEQKRAEA-----VAAMVWASCPNHPHSRVSTADNIYVNHSEFSGI-KG 1317
Db 1260 PVILSYGPDVNPALALEQKRAEA-----VAAMVWASCPNHPHSRVSTADNIYVNHSEFSGI-KG 1317
Qy 1318 AGAISNLFNPNRQF 1332
Db 1317 ANAARSLPKSQKF 1331

RESULT 12
US-10-663-208A-2
; Sequence 2, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442

```
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-663-208A-2

Query Match      78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 59

QY 61 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVAYEAFYQHSFABQSYDCSRVRVPA 179
Db 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVAYEAFYQHSFABQSYDCSRVRVPA 179

QY 180 RAATLAVGTMCVGYGSAALCNAQWLNFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEG 239
Db 180 RAATLAVGTMCVGYGSAALCNAQWLNFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEG 239

QY 240 VARNCSQDDVATCSQDCAASCPAIAAPQALDSTFYLGMPGSLVLIILCSVFAVVT 299
Db 240 VARNCSQDDVATCSQDCAASCPAIAAPQALDSTFYLGMPGSLVLIILCSVFAVVT 299

QY 300 ILLVGRVAPARDKSMVDPKGTSLSDKLSPTHTLLGQFQGGTWWASWPLTILVLS 359
Db 300 ILLVGRVAPARDKSMVDPKGTSLSDKLSPTHTLLGQFQGGTWWASWPLTILVLS 359

QY 360 VTPVVALAAGLVFTLTTDPVNLKAPNSQARSEKAFDQHEGPGFPFRNTQVLLTAPNRSS 419
Db 360 VTPVVALAAGLVFTLTTDPVNLKAPNSQARSEKAFDQHEGPGFPFRNTQVLLTAPNRSS 419

QY 420 VRYDSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSQDICVAPLNPDN 479
Db 420 VRYDSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSQDICVAPLNPDN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTLA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTLA 539

QY 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLAQAKLWEAFLE 599
Db 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLAQAKLWEAFLE 599

QY 600 NRAFQRMAGMFQVFTFAERSLEDRINRTADLPATSVIYIFLYISLALGSYSMSR 659
Db 600 NRAFQRMAGMFQVFTFAERSLEDRINRTADLPATSVIYIFLYISLALGSYSMSR 659

QY 660 VVWDSKATLGLGGVAVLGVANMAAGFFSYLGIRSLVILQVVPFLVSVGADNIFIVL 719
Db 660 VVWDSKATLGLGGVAVLGVANMAAGFFSYLGIRSLVILQVVPFLVSVGADNIFIVL 719

QY 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFGALTPMPAVRTFALTSLGA 779
Db 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFGALTPMPAVRTFALTSLGA 779

QY 780 VILDFLLQMSAFVALLSDSKRQESRLDVCCVKPQELPPQCGEGILLGFFQKAYAPF 839
Db 780 VILDFLLQMSAFVALLSDSKRQESRLDVCCVKPQELPPQCGEGILLGFFQKAYAPF 839

QY 840 LLHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFPLNRYPRVGAP 899
Db 840 LLHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFPLNRYPRVGAP 899

QY 900 VYFVVTLLGYNPSSEAGMNAICSSAGCNMPSFTQKIQTATEPPEQSIVLAIPASSWVDDFID 959
Db 900 VYFVVTLLGYNPSSEAGMNAICSSAGCNMPSFTQKIQTATEPPEQSIVLAIPASSWVDDFID 959

QY 960 WLTP-SSCCRLYISGPNKDKFCPSVNSLNCUKNCMSITMGSVRPSVRPQPHKYLWFLND 1018
Db 960 WLTP-SSCCRLYISGPNKDKFCPSVNSLNCUKNCMSITMGSVRPSVRPQPHKYLWFLND 1018

QY 1019 RNIKCPKGGGLAAYSTSVNLTSQGVLASRPMAYHKPKNSQDYTEALRAARELANITA 1078
Db 1020 RNIKCPKGGGLAAYSTSVNLTSQGVLASRPMAYHKPKNSQDYTEALRAARELANITA 1079

QY 1079 DLKRVPGTDPAPPEVFPYTHITVYEQYLTILPEGLPMLSLCLVPTTFAVNSCLLGLDLRSG 1138
Db 1080 DLKRVPGTDPAPPEVFPYTHITVYEQYLTILPEGLPMLSLCLVPTTFAVNSCLLGLDLRSG 1139

QY 1139 LNLLSIVMLIVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAIKTKPTWL 1198
Db 1140 LNLLSIVMLIVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAIKTKPTWL 1199

QY 1199 ERAKATISMSGSAVAGVAMTNLPGILVLGLAKAQLIQIIFPFLNLLITLGLLHGLVFL 1258
Db 1200 ERAKATISMSGSAVAGVAMTNLPGILVLGLAKAQLIQIIFPFLNLLITLGLLHGLVFL 1259

QY 1259 PVULSYVGPVNPALALQKRAEBAVAAVWASCPNHPSRVSTADNIYVNHSPESGI-KG 1317
Db 1260 PVULSYVGPVNPALALQKRAEBAVAAVWASCPNHPSRVSTADNIYVNHSPESGI-KG 1317

QY 1318 AGAISNLFNPNRGOF 1332
Db 1317 ANAARSSLPKSDQKP 1331

RESULT 13
US-10-646-301A-2
; Sequence 2, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-KI-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-646-301A-2

Query Match      78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 59

QY 61 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVAYEAFYQHSFABQSYDCSRVRVPA 179
Db 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVAYEAFYQHSFABQSYDCSRVRVPA 179

QY 180 RAATLAVGTMCVGYGSAALCNAQWLNFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEG 239
Db 180 RAATLAVGTMCVGYGSAALCNAQWLNFGQDTGNGLAPLDITFHLBPQAVGSGIQPLNEG 239

QY 240 VARNCSQDDVATCSQDCAASCPAIAAPQALDSTFYLGMPGSLVLIILCSVFAVVT 299
Db 240 VARNCSQDDVATCSQDCAASCPAIAAPQALDSTFYLGMPGSLVLIILCSVFAVVT 299

QY 300 ILLVGRVAPARDKSMVDPKGTSLSDKLSPTHTLLGQFQGGTWWASWPLTILVLS 359
Db 300 ILLVGRVAPARDKSMVDPKGTSLSDKLSPTHTLLGQFQGGTWWASWPLTILVLS 359

QY 360 VTPVVALAAGLVFTLTTDPVNLKAPNSQARSEKAFDQHEGPGFPFRNTQVLLTAPNRSS 419
Db 360 VTPVVALAAGLVFTLTTDPVNLKAPNSQARSEKAFDQHEGPGFPFRNTQVLLTAPNRSS 419

QY 420 VRYDSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSQDICVAPLNPDN 479
Db 420 VRYDSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSQDICVAPLNPDN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTLA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTLA 539

QY 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLAQAKLWEAFLE 599
Db 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLAQAKLWEAFLE 599

QY 600 NRAFQRMAGMFQVFTFAERSLEDRINRTADLPATSVIYIFLYISLALGSYSMSR 659
Db 600 NRAFQRMAGMFQVFTFAERSLEDRINRTADLPATSVIYIFLYISLALGSYSMSR 659

QY 660 VVWDSKATLGLGGVAVLGVANMAAGFFSYLGIRSLVILQVVPFLVSVGADNIFIVL 719
Db 660 VVWDSKATLGLGGVAVLGVANMAAGFFSYLGIRSLVILQVVPFLVSVGADNIFIVL 719

QY 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFGALTPMPAVRTFALTSLGA 779
Db 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFGALTPMPAVRTFALTSLGA 779

QY 780 VILDFLLQMSAFVALLSDSKRQESRLDVCCVKPQELPPQCGEGILLGFFQKAYAPF 839
Db 780 VILDFLLQMSAFVALLSDSKRQESRLDVCCVKPQELPPQCGEGILLGFFQKAYAPF 839
```

QY	120	FVNLHCHNTCSPNQSIFINVTVAQIAGAGQAPVAVYEAFFYQHFSASQSYDSCSRVRPA	179
DB	120	FVSLHCHNTCSPDQSLFINVTVRGAGEPPAVVAYEAFYQRFPAKAYECSQVRIPA	179
QY	180	AATLAVGTMCVYGSALCNAQRLWLFQDGTGNGLAPLDTITFHLLEPGQAVGSGIQIPLNEG	239
DB	180	AASLAVGSMCVYGSALCNAQRLWLFQDGTGNGLAPLDTITFHLLEPGQALPDGIQIPLNGK	239
QY	240	VARCNSQGGDDVATCSCODCAASCPAIRPQALDSTFYLGQMPGSLVLIILCSVPFVVT	299
DB	240	IAPCNSQGGDDSAVCSQDCAASCPVIPPPEALRPSFYMGMRPGWMLALIIIFTAVFVLLS	299
QY	300	ILLVGRFVAPARDKSKMVPDKGTSLSDKLSFSTHTLLGQFPQGGTGWASWPLTILVLS	359
DB	300	AVLVRVLRVSNRNKNAEGQEPAPKLPKHKLSPHTILGRPFQNGTRKVASWPLTVLALS	359
QY	360	VIPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQVILTPARNSS	419
DB	360	FIVVIALAAGLTIELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQIFVTPARNSS	419
QY	420	YRYDSLLGPKNFSGILDLLELELEQLERLHQLVMSPEAQRNISQDICYAPLAPNDN	479
DB	420	YKYDSLLGSKNFSGILDLLELELEQLERLHQLVMSPEAQRNISQDICYAPLAPYN	479
QY	480	TSLYDCINSLLQYFQNNRTLLLTANQITLMQGTQSDVMKQHFYLCANAPLTFKDGXTALA	539
DB	480	TSLSDDCVNSLLQYFQNNRTLLLTANQITLMQGTSLVDMKQHFYLCANAPLTFKDGXTSLA	539
QY	540	LSCMADYGAPVFPFLAIGVKGKDYGEAALTWTFSLNYPAGDPRLAQAKLWBEAFLEE	599
DB	540	LSCMADYGAPVFPFLAVGGVGGDYGEAALITFSLNYPADDPMAQAKLWBEAFLEE	599
QY	600	MRAFQRMAGMFQVTFTAERSLEDEINRTTAEDLPFATSYIVIFLYXLSLALGYSVSWR	659
DB	600	MESFORNTSKFQVAFSAERSLEDEINRTIQDLPVFAVSYYIVELYXLSLALGYSYRCSR	659
QY	660	VMDSKATLGLGVAVVGLVMAAMGFFSYLGRSLVILQVVPFLVLSVGADNIFIVL	719
DB	660	VAVESKATLGLGVVIVGLVAVLAAMGFYSYLVGPSSLVIIQVVPFLVAVGADNIFIVL	719
QY	720	EYQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTPMPAVRTFALTSGLA	779
DB	720	EYQRLPRMPGEPREAHIGRTLGSVAPSMLLCSLSEALCFPLGALTPMPAVRTFALTSGLA	779
QY	780	VILDFLLQMSAFVALLSLSKREASRLDYCCCVKQPELPPPCQGGELLGLFPQKAYAPP	839
DB	780	IILDFLLQMTAFVALLSLSKREASRPDLCCFSTRKLPPPKEKEGLLRFPRKIYAPP	839
QY	840	LLHWITRGVULLLPLAFGLVSLYSMCHISVGLDQELAPKDSYLLDYFLFINRYEVCAP	899
DB	840	LLHFIRIPVVMLLFLTFGANLYLMCNINVLGDLQELAPKDSYLLIDYFLFLNRYLEVGP	899
QY	900	VYFVTTLGYNPFSSSAGNATCSSAGCNFSFTOKIOVATEFPEOSYLAIPASSWDDFD	959
DB	900	VYFVTTSGFNPFSSSAGNATCSSAGCKSFSLTKIQVASEFPDQSVTAASWDDFD	959
QY	960	WLTP-SSCCRLYISGPNKDKPSTVSNLCLNCKMSITWGSVRPSVEQFHLYLPWFLMD	1018
DB	960	WLTPSSSSCCRLYIRGPHKDFECPSTDTSPFNCLNCKMNRITLGPVRPTAEQFHLYLPWFLMD	1019
QY	1019	RPNTKCPKGLAAVYSTVNLTSQDVLAISPMAYHKLKNSQDYITALAARELANITA	1078
DB	1020	PENIRCPKGLAAVYSTVNLSSDQVIAOSFMAHYHKLPLRNSODFTALRASLLANITA	1079
QY	1079	DLRKVPGTDPARFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSLIIGLDLRSG	1138
DB	1080	DLRKVPGTDPNFEVFPYITSNVYFQQYLTVLPEGIFTALCFVPTFVVCYLLGLDKCSG	1139
QY	1139	LINLLSIWMLVDITVGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSTKPTWL	1198
DB	1140	LINLLSIWMLVDITGLMAVNGISYNAVSLINLVAAGMSVEFVSHITRSTKPTRL	1199
QY	1199	ERAKEATISMGSAVFAGVAMTNLPGLIUVLGLAKAQLIQIIFFRNLNLTITLGLHGLVFL	1256
DB	1200	ERAKDATVFMGSAVFAGVAMTNFPGLIUVLGLFAQALIQIIFFRNLNLTITLGLHGLVFL	1259
QY	1259	PVILSYGCPDVNPALAJEQKAEBAVAAMVVASCPNHPSPRSTADNLYNVHSPGSI-KG	1317
DB	1260	PVILSYGCPDVNPALAJEQKAEBAVAAMVVASCPNHPSPRSTADNLYNVHSPGSI-KG	1317
QY	1318	AGATISNFPNNGROF	1332
DB	1317	ANARSSLPKSDQKF	1331
RESULT 14			
US-10-736-769-2			
; Sequence 2, Application US/10736769			
; Publication No. US20040161838A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Luquan			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JBO1603-K3-US			
; CURRENT APPLICATION NUMBER: US/10736.769			
; CURRENT FILING DATE: 2003-12-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; PRIOR APPLICATION NUMBER: 10/663,208			
; PRIOR FILING DATE: 2003-09-16			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 2			
; LENGTH: 1331			
; TYPE: PRT			
; ORGANISM: Rattus sp.			
US-10-736-769-2			
Query Match 78.5%; Score 5421.5; DB 4; Length 1331;			
Best Local Similarity 77.8%; Pred. No. 0;			
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6			
QY	1	MABAGURGMWLLWALLRLUAQSEPYTHIQPGYCAFYDCGKNGPELSGMTLSNVCSLGN	60
DB	1	MAAAWL-GWLLWALLLSAAQGELYTPKHEAGVCTFYECGKNGPELSGMTLSNVCSLGN	59
QY	61	TPARKITGDHLLILLOKICPRLVTGNPTQ-ACCSAQLVSLSEASLITKALLTRCPACSDN	119
DB	60	TPARHVTGHEALLQRIQCPRLNGNPTTFACCTQKLSLESMSITKALLTRCPACSDN	119
QY	120	FVNLHCHNTCSPNQSIFINVTVAQIAGAGQAPVAVYEAFFYQHFSASQSYDSCSRVRPA	179
DB	120	FVSLHCHNTCSPDQSLFINVTVRGAGEPPAVVAYEAFYQRFPAKAYECSQVRIPA	179
QY	180	AATLAVGTMCVYGSALCNAQRLWLFQDGTGNGLAPLDTITFHLLEPGQAVGSGIQIPLNEG	239
DB	180	AASLAVGSMCVYGSALCNAQRLWLFQDGTGNGLAPLDTITFHLLEPGQALPDGIQIPLNGK	239
QY	240	VARCNSQGGDDVATCSCODCAASCPAIRPQALDSTFYLGQMPGSLVLIILCSVPFVVT	299
DB	240	IAPCNSQGGDDSAVCSQDCAASCPVIPPPEALRPSFYMGMRPGWMLALIIIFTAVFVLLS	299
QY	300	ILLVGRFVAPARDKSKMVPDKGTSLSDKLSFSTHTLLGQFPQGGTGWASWPLTILVLS	359
DB	300	AVLVRVLRVSNRNKNAEGQEPAPKLPKHKLSPHTILGRPFQNGTRKVASWPLTVLALS	359
QY	360	VIPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQVILTPARNSS	419
DB	360	FIVVIALAAGLTIELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQIFVTPARNSS	419

Db	783	DFLLQMTAFVALLSLDSKROEASRPDVVCCFSSRNLPKPKQEGLLCCFRIKYTFPFLH	842
Qy	843	WIRGVULLFLALFGVSLYSMCHISVGLDOELALPKDSVLLDYDFLNLRYEVEGAPVVF	902
Db	843	RFIRPVULLFLVLFGANLYLMCNISVGLDOELALPKDSVLLDYDFLNLRYEVEGPPVVF	902
Qy	903	VTTLYGNFSSEAGNNAICSSAGCWNFSFTQKIQYATEFFPEQSYLAIPASSWVDDFDWLT	962
Db	903	DTTSGYNFSTEAGNNAICSSAGCESFSLTQKIQYASEFPNQSVTAIAASSWVDDFDWLT	962
Qy	963	P-SSCCRLYISGPNKOKPCPSTVNSLCLKNCMSITWGSVRPSVEQPHKYLFWFLNDRPN	1021
Db	963	PSSCCRIYTRGPHKDFCPSDTDTSPCLKNCMNRTLGPVRTTEQPHKYLFWFLNDRPN	1022
Qy	1022	IKCPKGGLAAYSTSVNLTSDGOVLASFPMAYHKPLKNSQDYTALRAARELANITADLR	1081
Db	1023	IRCPKGGLAAYRTSVNLTSSDQIIITASQFMAYHKPLRNSQDFTALRASRLAANITAEUR	1082
Qy	1082	KVPGTDPAFEVPPTITNVFVEQVLTTLPGELFMLSCLVPTFAVSLGLGLDLRSGLLN	1141
Db	1083	KVPGTDNFEVPPTISNVFYQQYTLVLPGEIFTLALCFVPTFVVCYLLGLDLRSGIILN	1142
Qy	1142	LLSYIMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER	1201
Db	1143	LLSYIMILVDIIGLMAVWGISYNAVSLINLVTAGMSVEFVSHITRSFAVSTKPTLER	1202
Qy	1202	KEATISMGSAVFAVAMTNLPGIILVGLAKAQLIOIFFRNLNLITLLGLLHGLVFLPVI	1261
Db	1203	KDATIFMGSAVFAVAMTNFPGILILGPAQQLIOIFFRNLNLITLLGLLHGLVFLPVV	1262
Qy	1262	LSYVGPDVNPALAEQKRAEBAVAMVAVSCPNIHPSRVSTADNITVNVHSPGS-IGGAGA	1320
Db	1263	LSYLGPDVNQALVLEEKLEATEA-AMVSEPSQCFPPFADANTSDEVYNGNPEFPEINA	1321
Qy	1321	ISNFLPNNGROF	1332
Db	1322	ASSSLPKSDQKF	1333

Search completed: April 7, 2006, 19:37:11
Job time : 194 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:34:15 ; Search time 27 Seconds
(without alignments)
1538.804 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGISNPLPNRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 3119182 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

Database :

Published Applications AA_New:*
1: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /SIDSS/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /SIDSS/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDSS/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
6: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /SIDSS/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6896	99.8	1332	7	US-11-242-459-9
2	2402.5	34.8	1278	6	US-10-995-561-952
3	942	13.6	194	7	US-11-242-459-8
4	195.5	2.8	1087	7	US-11-242-459-17
5	195.5	2.8	1392	7	US-11-242-459-40
6	191	2.8	127	7	US-11-096-568A-6778
7	173	2.5	85	7	US-11-096-568A-6779
8	143	2.1	445	7	US-11-242-459-35
9	143	2.1	456	7	US-11-242-459-34
10	141.5	2.0	1048	6	US-10-392-234A-18
11	140.5	2.0	1046	6	US-10-392-234A-16
12	134	1.9	1067	6	US-10-467-657-2548
13	127	1.8	1048	6	US-10-392-234A-14
14	126.5	1.8	200	7	US-11-242-459-16
15	126	1.8	801	6	US-10-793-626-2020
16	121.5	1.8	1043	6	US-10-392-234A-34
17	120.5	1.7	2426	7	US-11-203-806A-11
18	119	1.7	1048	6	US-10-392-234A-20
19	117.5	1.7	1049	6	US-10-392-234A-12
20	115.5	1.7	914	7	US-11-072-512-2923
21	113	1.6	1034	6	US-10-392-234A-30
22	112.5	1.6	697	7	US-11-082-389-362
23	109.5	1.6	758	7	US-11-096-568A-30412
24	109.5	1.6	772	7	US-11-096-568A-30411
25	109.5	1.6	827	7	US-11-096-568A-30410

26	109.5	1.6	1049	7	US-11-098-686-11239	Sequence 11239, A
27	109	1.6	1032	6	US-10-392-234A-67	Sequence 67, Appl
28	108.5	1.6	648	7	US-11-087-099-4135	Sequence 4135, Ap
29	106	1.5	443	7	US-11-087-099-11457	Sequence 11457, A
30	105.5	1.5	783	7	US-11-082-389-354	Sequence 354, App
31	105	1.5	324	7	US-11-087-099-11792	Sequence 11792, A
32	105	1.5	1080	7	US-11-096-568A-27723	Sequence 27723, A
33	105	1.5	1097	7	US-11-096-568A-27722	Sequence 27722, A
34	105	1.5	1128	7	US-11-096-568A-27721	Sequence 27721, A
35	104.5	1.5	315	7	US-11-087-099-9820	Sequence 9820, Ap
36	104.5	1.5	471	7	US-11-087-099-4362	Sequence 4362, Ap
37	104	1.5	757	7	US-11-096-568A-27509	Sequence 27509, A
38	104	1.5	771	7	US-11-096-568A-27508	Sequence 27508, A
39	104	1.5	826	7	US-11-096-568A-27507	Sequence 27507, A
40	103.5	1.5	251	7	US-11-087-099-1098	Sequence 1098, Ap
41	103.5	1.5	331	7	US-11-087-099-8831	Sequence 8831, Ap
42	103.5	1.5	2042	6	US-10-506-454-134	Sequence 134, Appl
43	103.5	1.5	2312	7	US-11-126-313-34	Sequence 34, Appl
44	103	1.5	463	7	US-11-087-099-3980	Sequence 3980, Ap
45	102.5	1.5	475	7	US-11-087-099-2723	Sequence 2723, Ap

ALIGNMENTS

RESULT 1

US-11-242-459-9
; Sequence 9, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-9

Query Match 99.8%; Score 6896; DB 7; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60	
Db	1	MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60	
Qy	61	TPARKITGDHLILLOKICPRLYTGTNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120	
Db	61	TPARKITGDHLILLOKICPRLYTGTNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120	
Qy	121	VNLHCHNTCSNPQSLFINVTRVAQIGAGOLPAVVAEYAFYQHSFAEQSDVSCSRVRPAA 180	
Db	121	VNLHCHNTCSNPQSLFINVTRVAQIGAGOLPAVVAEYAFYQHSFAEQSDVSCSRVRPAA 180	
Qy	181	ATLAVGTMCYGVYGSALCNQAWLNFGQDTGNGLAFLDITFHLLRPGQAVGSIQPLNEGV 240	
Db	181	ATLAVGTMCYGVYGSALCNQAWLNFGQDTGNGLAFLDITFHLLRPGQAVGSIQPLNEGV 240	
Qy	241	ARCNSQGDVATCSQDCAASCAPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300	
Db	241	ARCNSQGDVATCSQDCAASCAPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300	

301 QY LLVGRVAPARDKSMVDPKKGTSLSKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360
301 Db LLVGRVAPARDKSMVDPKKGTSLSKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360
361 QY IPVVALAAGLVFTLTTDPVBLWSAPNSQARSEKAFHDQHFGPPFRTNQVILTAPNRSY 420
361 Db IPVVALAAGLVFTLTTDPVBLWSAPNSQARSEKAFHDQHFGPPFRTNQVILTAPNRSY 420
421 QY RYDSLGLGPKNFGSLDLDLLELELERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
421 Db RYDSLGLGPKNFGSLDLDLLELELERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
481 QY SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGATLAL 540
481 Db SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGATLAL 540
541 QY SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
541 Db SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
601 QY RAQORMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
601 Db RAQORMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
661 QY MYDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVSVGADNIPFVLE 720
661 Db MYDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVSVGADNIPFVLE 720
721 QY YQRLPRRPGEPREVIHGRALGVAPSMLLCSLSEACPFELGALTMPAVRTEALTSLAV 780
721 Db YQRLPRRPGEPREVIHGRALGVAPSMLLCSLSEACPFELGALTMPAVRTEALTSLAV 780
781 QY ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQEGLLGFPQKAYAPPL 840
781 Db ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQEGLLGFPQKAYAPPL 840
841 QY LHMTRGVVLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFLNRYFEVGA 900
841 Db LHMTRGVVLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFLNRYFEVGA 900
901 QY YFVTTLYGNPSSSAGNNAICSSAGCNPFSTQKIQYATFPPQSYLAIIPASSWDDFIDW 960
901 Db YFVTTLYGNPSSSAGNNAICSSAGCNPFSTQKIQYATFPPQSYLAIIPASSWDDFIDW 960
961 QY LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQFHKYLPWFNDLRP 1020
961 Db LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQFHKYLPWFNDLRP 1020
1021 QY NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPKLNQSDYTEALRAARELAANITADL 1080
1021 Db NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPKLNQSDYTEALRAARELAANITADL 1080
1081 QY RKVPGTDPAFEPVFTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
1081 Db RKVPGTDPAFEPVFTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
1141 QY NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
1141 Db NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
1201 QY AKENATISMGSAVPAGVANTMLPGLVLGLAKAQLIQIFFRLNLLITLGLLHGLVFLPV 1260
1201 Db AKENATISMGSAVPAGVANTMLPGLVLGLAKAQLIQIFFRLNLLITLGLLHGLVFLPV 1260
1261 QY ILSVVGDPVNPALAEQKRAEEAVAAVMVASCNHPRSVSTADNIYVNHSEFSGIKGAGA 1320
1261 Db ILSVVGDPVNPALAEQKRAEEAVAAVMVASCNHPRSVSTADNIYVNHSEFSGIKGAGA 1320
1321 QY ISNFLPNNGRQF 1332
1321 Db ISNFLPNNGRQF 1332

RESULT 2
US-10-995-561-952
; Sequence 952, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 952
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-952

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.6e-181;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGMILLMALLRLAQSEPVTTIHQPGYCAFYDECG-----KNPELSGSLMTLSNVVCLS 59
Db 4 RGLALGULLLLCPAQVFSQ-----SCWYGECCGAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLILLOKICPRLYTGPNTPQACCSAKQLVSLBASISYTKALLTRCPACSDN 119
Db 47 --PKPLPKDGYDLVQELCPGFFFG--NVSLCCDVRQLQTLKDNLQLPLQFSLRSCPCFYN 103
QY 120 FVNLHCHTCSPNOSLFINVTR-----VAQLGAGQLPAVAYEAFVQHSFASQSDSCSRV 175
Db 104 LNLFLCELTCSPROSQFLNVLTATEDYVDPVTNQTKTNVKELOYVYVGQSFANAMYNACRDV 163
QY 176 RVPAATLAVGTMCVYGSALCNAQRWLNFOGDTGNGLAPLDIT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA--CNATNWEYFMKNDGQAPFTITVPVDFPVH----- 215
QY 229 VGSQIQLNEGVARCNESQGDVATCSQDCQCAASCPAARFQ-----ALDSTFYLG 279
Db 216 --GMEPNNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPAPPTILGLDAMYVIM 272
QY 280 QMPGSLVLIILCSVPFVVITLL-----VGRVAPARDKSMVDPKKGTSLS 326
Db 273 WITYMAFLLLPFGAFVAVWCYRKRKYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFQGGTGWASWPLTILVLSVIVVVALAAGLVFTLTTDPVBLWSAP 386
Db 325 DPVSAAFEGCLRRLFTRWGSFCVRNPGCVIPFSLVFTACSSGLVFRVVTNPVBLWSAP 384
QY 387 NSQARSEKAFHDQHFGPPFRTNQVILTAPNRSYSDLSLLGPK--NFSGLDLDLLELL 445
Db 385 SSQARLEKEYPDQHFQFPFRTQLIRAPLTDKHIYQYPYGADVPFGPPLDIQILHQL 444
QY 446 ELQERLHLQVMSPEAQRNLSLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLTFA 505
Db 445 DLQIAIEN--ITASVDNETVTIQLDCLAPLSPYNT---NCTILSVLNYFQNSHVLDDHK 499
QY 506 NOTLMGQTSQVMDKDHFLYCANAPLTFKDGATLALSCMADYGAPVFPFLAIGGYKGDYS 565
Db 500 GDDFF---VYADVHTFLYCVRAPASLNDTSLDHPCLGTGPGVFPVFWLVLGGYDQNYN 556
QY 566 EBAALIMTFSLNYPAGDPRLAQAKLWEEAFLEEMBAFORWAGMFOVTTFAERSLEDEI 625
Db 557 NATALVITFPVNYNDTEKLRQAQAEKSEFINVKNYKN---PNLITISFTAERSLEDEL 613
QY 626 NRTTAEDLPFATSYIVIFLYISLALGSYSSWSRVSMVDSKATLGLGGVAVVLGAVMAAG 695
Db 614 NRESDSQVTVVISYAIMFLYISLALGHKSCRLLVDSKSVLSGIAGILIVLSSVACSLG 673

QY 686 FFSYLGRSSVLQVVPFLVSVGADNIFIFVLEYQRLPRRPGEPREHIGRALGRVAP 745
Db 674 VFSYIGLPTLVIVIEPVLAVGVNDNFILVQAYQDERLQGETDQDLGRVLGEVAP 733
QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSDSKRQAS 805
Db 734 SMFLSSFSVTAFFLGALSMVPAVHTFSLPAGLAVFIDFLLAITCFVSLGLDIKRQKN 793
QY 806 RLDDVCCVPQBLPPGQ-GEGLLGGFFOKAVAPFLLHWITRGVLLLFALFGVSLYSM 864
Db 794 RLDFCCVRGAEDGTSVQASESCLFFPKNSYSPILLKDWMPVIAIFVGVLSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDFLNRVFEVQAPVYFTTILGYNPSSAGNAICSSAG 924
Db 854 NKVDIGLQSLSPDDSYNVDYFKSISQYLHAGPPVYFVLEBHDYTSKGGQNMVCGG 913
QY 925 CNFSTQKIQVATPEPQSYLAIPASSWVDDFIDWLTP-SSCCRLYISGPNKDKPCPT 983
Db 914 CNNDLSVQOIFNAQOLDNVTIRGFAPSSWIDDYFDWVKQSSCCRV---DNITDQFCNAS 970
QY 984 VNSMLCNKCMSIT-MGSVRPSVEQPHKYLPWFLNDRPNKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRLTPBQKQPGQDFMFLPFLSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQCVLASRPMAYHKLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNV 1100
Db 1030 GTRVGATVEMTYHTVLTQTSADPIDALKKARLIASNT-ETMGINGS--AYRVFPYSVFV 1086
QY 1101 FVEQYLTILPEGLFMLSLCLVPTFAVSCLLGLDLSRGLNLLSVMLIADTVGPMALWD 1160
Db 1087 FVEQYLTIDDDIFNLGSLGAIFLVTWVLLGCELWSAVIMCATAMVLVNMFGVWMLG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSPAIKPTWLERAKEATISMSAVFAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGSGSVEFCSHITRAFTVSMKGSRVERAEALAHMGSSVFSGITLK 1206
QY 1221 LPGAIVLGLAKAQLIQIFPRLNLLITLGLLHGLVFLPVLISYVGPDPNPA 1272
Db 1207 FGGIVVLAPAKSQIQIFFRMYLAMVLLGATHGLIFLPLVLSYIGPSYKA 1258

RESULT 3

US-11-242-459-8
; Sequence 8, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 8
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-8

Query Match 13.6%; Score 942; DB 7; Length 194;
Best Local Similarity 99.5%; Pred. No. 5.1e-67;
Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 615 FTAERLEDEINRTTAEDLPATATSYIVIFLYISALGYSWSRWVDSKATLGLGVA 674
Db 1 FMAERLEDEINRTTAEDLPATATSYIVIFLYISALGYSWSRWVDSKATLGLGVA 60

QY 675 VVLGAVMAAMGPFSSYLGRSSVLQVVPFLVSVGADNIFIFVLEYQRLPRRPGEPREV 734
Db 61 VVLGAVMAAMGPFSSYLGRSSVLQVVPFLVSVGADNIFIFVLEYQRLPRRPGEPREV 120
QY 735 HIGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVAL 794
Db 121 HIGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVAL 180
QY 795 LSLDSKROEASRLD 808
Db 181 LSLDSKROEASRLD 194

RESULT 4

US-11-242-459-17
; Sequence 17, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-17

Query Match 2.8%; Score 195.5; DB 7; Length 1087;

Best Local Similarity 19.5%; Pred. No. 6.1e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEENRAFQORMAGMFQVFTTFAERSLED----- 623
Db 100 LPNTYSVDDR-----WEE-----QRAKFSQSVTVYVAVLAKQSTSKVQVLYGGTD 144
QY 624 ----EINRTTAD--LPIPATSYIVIFLYISALGYSWSRWVDSKATLGLGG-VAVV 676
Db 145 LFDYEVRETRTFNNDMLLAFISSCIAALVYILTSCSVFLSFFGI-----ASIGLSCLVALP 199
QY 677 LGAVMAAMGPFSSYLGRSSVLQVVPFLVSVGADNIFIFVLEYQRLPR-RPGEPREH 735
Db 200 LYHVVFGEI---OYLGILNG-----VAAPVIGIGVDDVFPVFTNTYRQATHLEDQPLRMTH 251
QY 736 IGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVI---LDLFLQMSAFV 792
Db 252 TVQTAGK---ATPFTSLTTAAAYANVFSQIPAVHDGFLPMSLIVSCCWLAVLVTMPAL 308
QY 793 ALLSLDSKROEASRLDVC--CCVKPQBLPPGQ----- 823
Db 309 GLWSLYLAPLESSCQTSCHQNSRKTSLHFPDGDVFAAPEQVGGSPAQGPPIPYLDDDDIPL 368
QY 824 -----GEGILLGFFOKAYAP-----FLAHWI-----TRGVVL 850
Db 369 EVEEPVSELEIGDVSLSVSPGELQPNASNTSGRHLIVOLQELHLLHHWMSAVKSRWIV 428
QY 851 LFLALFGVSL-----YSM-----CHISVGLDQ- 874
Db 429 GLFVSILLILSVFASRLRPAAPLLFRPDNIQVLLDKYNLSAEGISCITCSGLPQEK 488
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFVGAPVFPVFTTILGYNF 910
Db 1 FMAERLEDEINRTTAEDLPATATSYIVIFLYISALGYSWSRWVDSKATLGLGVA 60

Db	489	PHSLQNNIRTSLEKRRGSGVFWASRPENTLQDPFGTVYISKVKSQGHQA--VYRLSLNA	546
Qy	911	SSEAGMNAICSSAG--CNMF-----SFTOKIQVATEFPBQSLALIPASSWVDDFDWL	961
Db	547	SLPAPMQAVSPGDGEVPSFQVYRAPFGNFTKLTACHMSTVGLLQAASPSRKM-----L	600
Qy	962	TPSSC-----CRLYIS--GP	974
Db	601	TTLACDAKRGWKDFSFYVATKEQOHTKLYFAQSHKPPPHGRVCMAPPGLSSSPDG	660
Qy	975	NKDFCPSVNSLNCNKMISITWG-----SVRPSVEQPHKLPWF-----LN-	1017
Db	661	TKGFF---FVPSEKVPKARLSATGFNCPVNTGCGKPAVRLVDTGAMVFFVGLIGVNR	717
Qy	1018	---DRPNIKCPKGLAAAYSTSVNLTS-----GOVL	1045
Db	718	TRQVDNHHVIGDP--GSVYDSSFDLFKEIGHLCHLCKAIAANSELVKPGGAQCLPSGYSI	775
Qy	1046	ASRFMAYHKPLKNSQD-----YTEALRAARELANITADLRKVPGTDPAPFVP-P	1094
Db	776	SSFLQMLHPECKELPEPNLLPGQLSHGAVGVRGVRQVISMAFESTTYKGS--SFQYSD	834
Qy	1095	YTIITNVFEQYLTILPEG-----LFW-----LSLCL--VP	1122
Db	835	YLRWESFLQQLQALPEGSVLRGFGTCEHWKQIFMEIVGVQSALCGLVLSLLICVAAVA	894
Qy	1123	TFVAVSLGLDLRLSGLNLLSIVMLVDTVGFMAWDISYNAVSLINLVSAGVMSBVF	1182
Db	895	VFTTHILL-----LLPVLSILGIVCLV--TIMYWSGEM--GAVEAISLSILVGSVDYC	947
Qy	1183	SHITRSFAIS-----TKPTWLERAKEATISNGSAVFAGVAMTNLPGI-----	1224
Db	948	VHLVEGYLLAGENLPPHQAEDARTQW--RTLEAVRHVGVAIVSSALTVTIATVPLPFC	1005
Qy	1225	LVLGLAKAQIQLIFFPRMLNLLTLGLLHGLFPLVILSYVGPVNPALALEQKRAE--E	1282
Db	1006	IIAPPFAK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTRTTSFLK	1053
Qy	1283	AVAAVMVA	1290
Db	1054	ALGAVLLA	1061
RESULT 5			
US-11-242-459-40			
; Sequence 40, Application US/11242459			
; Publication No. US20060035835A1			
; GENERAL INFORMATION:			
; APPLICANT: TANIYAMA, Yoshio			
; APPLICANT: KITA, Shunbun			
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof			
; FILE REFERENCE: 2703USOP			
; CURRENT APPLICATION NUMBER: US/11/242,459			
; CURRENT FILING DATE: 2005-10-03			
; PRIOR APPLICATION NUMBER: US/10/239,316			
; PRIOR FILING DATE: 2002-09-19			
; PRIOR APPLICATION NUMBER: PCT/JP01/02279			
; PRIOR FILING DATE: 2001-03-22			
; PRIOR APPLICATION NUMBER: JP2000-088595			
; PRIOR FILING DATE: 2000-03-24			
; NUMBER OF SEQ ID NOS: 59			
; SEQ ID NO 40			
; LENGTH: 1392			
; TYPE: PRT			
; ORGANISM: Human			
US-11-242-459-40			
Query Match 2.8%; Score 195.5; DB 7; Length 1392;			
Best Local Similarity 19.5%; Pred. No. 8.6e-07;			
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;			
Qy	576	LNNYPAGDPRLAQAKLWEEAFLEWRFAFORMAGNFQVTFTAERSLED-----	623

US-11-096-568A-6778
; Sequence 6778, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6778
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(127)
; OTHER INFORMATION: Ceres Seq. ID no. 15168324
US-11-096-568A-6778

Query Match 2.8%; Score 191; DB 7; Length 127;
Best Local Similarity 44.3%; Pred. No. 6.4e-08;
Matches 35; Conservative 21; Mismatches 23; Indels 0; Gaps 0;

QY 1199 ERAKERTISMGSAVAGVAMTNLPGILVGLAKAQLIQIFFRNLNLLITLGLHGLVFL 1258
Db 34 QRAKTALCTMGASVSGITLTKLVGLVLCFSTSIFFVYVYFQMYLALVIGLHGLVFL 93

QY 1259 PVILSVGPDVNPALAEQ 1277
Db 94 PVVLSLFGPPRYTVIKEQ 112

RESULT 7
US-11-096-568A-6779
; Sequence 6779, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6779
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(85)
; OTHER INFORMATION: Ceres Seq. ID no. 15168325
US-11-096-568A-6779

Query Match 2.5%; Score 173; DB 7; Length 85;
Best Local Similarity 44.3%; Pred. No. 9.6e-07;
Matches 31; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 1208 MGSVAGVAMTNLPGILVGLAKAQLIQIFFRNLNLLITLGLHGLVFLPVILSVGP 1267
Db 1 MGSVSGITLTKLVGLVLCFSTSIFFVYVYFQMYLALVIGLHGLVFLPVILSVLFGP 60

QY 1268 DVNPALAEQ 1277
Db 61 PLRYTVIKEQ 70

RESULT 8
US-11-242-459-35
; Sequence 35, Application US/11242459
; Publication No. US20060035835A1

US-11-096-568A-6778
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 35
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-35

Query Match 2.1%; Score 143; DB 7; Length 445;
Best Local Similarity 25.4%; Pred. No. 0.0024;
Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;

QY 576 LNNYPAGDPRLAQAKLWBEAFLEEMRAFORMAGMFQVFTTAERSLED----- 623
Db 100 LNNYSVDDR-----WEE-----QRAKFSQPVVTVVAMLAQKSTSKVQVLYGGTD 144

QY 624 ----BINTTARD--LPITATSYIVIFLYISIALGYSWSRVMVDSKATLGG-VAVV 676
Db 145 LPDYEVRRTFNDMLLAFITSSCIAALVYILTSCSVLSFFGI-----ASIGLSCLVALF 199

QY 677 LGAVMAAGPFSYLGIRSLVILQVVPFLVLSVGADNIFVLEVQRLPR-RPGEPRVH 735
Db 200 LYHVVFGI---QYLGILNG-----VAAFPVIGVIGVDVFEVFTYRQATHLEDPPQLRMTH 251

QY 736 IGRALGRVAPSMILGSLSEAICTFLGALTTPMPAVRTFALTSLGLAVI---LDFLLQMSAFV 792
Db 252 TVQTAGK---ATPFTSLTAAAYANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308

QY 793 ALLSLSKQREASRLDVC--CCVKPQELPPPG 822
Db 309 GLWSLYLAPLESQCTSCHQNCRTSLHFPG 340

RESULT 9
US-11-242-459-34
; Sequence 34, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 34
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-34

Query Match 2.1%; Score 143; DB 7; Length 456;
Best Local Similarity 25.4%; Pred. No. 0.0025;

Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;

QY 576 LNNTYAGDPRLAQAKLWEEAFLEENRAFQRRMAGMFQVFTFAERSLED----- 623
DB 100 LPNYYSVDDR-----WEE-----QRAKFSQVVVYVAMLAQKSTSKVQVLYGGTD 144

QY 624 ----EINTTAED--LPFATSYIVIFLYISLALGYSWSRVMVDSKATLGLGG-VANV 676
DB 145 LFDVEVRTFNNDMLLAFISSCIAALVYILTSCVFLSPFGI-----ASIGLSCLVALF 199

QY 677 LGAYMAAGFFSYLGIRSSLVLQVVPFLVSLVAGADNIFIFVLEYQRLPR-RPGEPREVH 735
DB 200 LYHVVFGI---QYLGILG-----VAAPVYIGVDDVFVINTYRQATHLEDQLRWIH 251

QY 736 IGRALGRVAPSMMLCSLSEAICFFLGAUTPMPAVRTFALTSLGLAVI---LDFLQMSAFV 792
DB 252 TVQTAGK---ATFTSLTAAAYANVFSQIPAVHDFGLFSLVSCCMCLAVLVTMPAAL 308

QY 793 ALLSDSKRQEARLDVC---CCVKQELPPPG 822
DB 309 GLWSLYLAPLESSCQTSCHQNCRSKRTSLHFFG 340

RESULT 10
US-10-392-234A-18
; Sequence 18, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxseer, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiaozhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-392-234A-18

Query Match 2.0%; Score 141.5; DB 6; Length 1048;
Best Local Similarity 18.1%; Pred. No. 0.011;
Matches 145; Conservative 110; Mismatches 243; Indels 303; Gaps 32;

QY 624 EINTTAEDLPFATSYIVIFLYISLALGYSWSRVMVDSKATL--GLGGVAVLGG--A 679
DB 339 EVVKTLVE-----AIIIVFLWYLFQNF-----RATLPTIAVPVLLGTEA 381

QY 680 VMAAMGFFSVLGRSSLVLQVVPFLVSLV--ADNIFVLEYQRLPRPGEPREVHIG 737
DB 382 VLAAGF-----SINTLTMFG--MVLAIGLLVDDAIVVENVERVMTTEEGLPKBEATR 432

QY 738 RALGR-----VAPSMMLCSLSEAICFFLGAUTPMPAVRTFALTSLGLAVIILDFLQMSAFV 792
DB 433 KSMGQIQGALVGIAMVLSAVFIPWAFGGSTGAI--YRQPSITIVSMALSVLVALILTP 490

QY 793 ALLSDSKRQEARLDVCCVKQELPPPGQSGLLIGFFQKAYAPFLHVI----- 844
DB 491 ALCA-----TMLKPVAKDGHGKKGFGWENFLFDKSTHYTDSVGNILR 536

QY 845 TRGVLLLLFLALF-----QVSLYSMCHISVGLDQLALPKDSYLL 884
DB 537 STGRYLLLYIIIVGMAYLFRVLPSFLPDEDQGVFL--TWQLPAGATQERTOKVLDEVT 595

QY 885 DYFL-----FLNR----- 892

DB 596 DYLYNKEKANVESVPAVNGFGFAGRCQNTGIAFVSLKDWADRPGEKKNKVEAITQRATAAF 655

QY 893 -----YFEVGAPVYFVTTLGYNFSSEAGNNAICSSAGCNCNFSQT----- 932
DB 656 SOIKDAMVAFNLPALVELGTATGDFPE-----LIDQGLGHEKUTQARNQLFGEVAK 708

QY 933 -----KIQVATEPEQSYLAIPAS-----SWVDDPFDWLTP 963
DB 709 YPDLLVGVRRNGLEDTPQFKIDIDQEKQAQALGVISIDINTTLGAAMWGSVYNDPID--R 765

QY 964 SSCRLYISGNPKDKFCPSVNSLNCNLCNMSITMGSVRSVQFHKYKLPFLNDRPNIK 1023
DB 766 GRVKKVVYVMSEAKYRMLPDDIND-----WYVRSQGQM 798

QY 1024 CPKGLAAYST-----SVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELANIT 1077
DB 799 VP---FSAFSSRWYSGSPRLRYNGLPSMEILGQAAPGKST---GEAMAMEELASKLP 852

QY 1078 ADLRKVPGTDPAPFVFPYITITNVFYEQLT-----ILPEGLFMLSCL----- 1120
DB 853 SGI-----CYDWTGMSYQERLSGNQAPALYALSLIVVFLCLAALYESWSIPF 899

QY 1121 ----VPTFAVSCLLGLDLRLSGLNLLSIVMILVDTVTFPMALWDISYNAVSLN----- 1170
DB 900 SVMLVPLGVIGALLAA--TFRGLTNDVYFQVGLLTITGLSA-----KNAILIIVEFAKDL 952

QY 1171 -----LVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMT 1219
DB 953 MDKEGKGLVEAMLEAVRMRLPILMTSLAPMLGVMP-----LVISSG-----AGSGAQ 1000

QY 1220 NLPGLVL--GLAKAQLIQIFP 1239
DB 1001 NAVGTGVLGVMVATVLAIFP 1021

RESULT 11
US-10-392-234A-16
; Sequence 16, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxseer, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiaozhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 1046
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-392-234A-16

Query Match 2.0%; Score 140.5; DB 6; Length 1046;
Best Local Similarity 20.5%; Pred. No. 0.013;
Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;

QY 506 NOTLMGOT--SQVDMKDHFLYCANAPLTPKDGCTALALSCWADYAGVFPFPLAIGGYKGD 563
DB 231 NATIIGTKRLQTAEPENILLKVN-----PDGSOVRLKDVADV-----LG---GQD 274

QY 564 YSEAEALIMTFLNNYPA-----GDPRLAQAKLWEE--AFLEEMRAFQRRMAGMFQ 612
DB 275 YS-----INAQFNGSPASGIAIKLATGANALDTAKAIRQTIANLEPFPQGMKVVPYD 328

QY 613 VPTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGYSWSRVMVDSKATL--GL 670

Db 329 TTPVVSASTH-EVVKTLGE-----AIIIVFLVMYLFLONF-----RATLIPTI 370
QY 671 GGVAVVLG--AVMAAMGFFSYLGIRSSVLQVVPFLVSVG--ADNIFIFVLEYQRLPR 726
Db 371 AVPVVLGTFGVLAAGF-----SINTLMFG-MVLAIGLLVDDAIVVVVENRVMA 421
QY 727 RQG-EPREV--HIGBALGR-VAPSMLLCSLSEAICFFLGALTPMPAVTTPALTSLGLAVI 781
Db 422 BEGLSPREARKSMQIQIAGVIAVLSAVFLPMAFFPGSGTVI--YRQFSITIVSAMA 479
QY 782 LDFLLQMSAFVALLSLDSKRQASRLDVCC--VKPQELPPPGEGGLLGFQKAYAP 838
Db 480 L-----SVIVALI-----LTPALCATMLKPIEKGDHGEHKGFFGFWNMFUS 522
QY 839 FLLHMITRGV-----LILLFALFGVSLYSMCHISVGL-----DQ-----ELALPKD 880
Db 523 -TTHGYERGASILKHRAFPYLLIYVIVVAGMIWMFTRIPTAFLPDDEQGVLAQVTPPG 581
QY 881 S-----YLLDYFLFLNRYFVGAPVPTVTLGNFSSB----- 913
Db 582 SSAERTQVVVDSMRREYLLEK-----ESSVSSVFTVTGTFAGRGQSSGMATMLKP 633
QY 914 ----AGWNAI--CSSACNNFSTQKIQYATERPEQSYLAIPASSWVDDFID----- 959
Db 634 WEERPGGENSVFELAKRAQMHFFSFKDAMVFA--FAPPSVLBNATGDFLFDQOAGVG 691
QY 960 ----WLTSSCCRLYISGNKDKFCPSTVN-----SLNCLKNCMSI 996
Db 692 HEVLQARNKFLMAAQNALQVRPNMGMSDEPQYKLEIDDEKASALGVSLADINSTVSI 751
QY 997 TWGSV-----RPSVEQPHKYLPLNDRPN-----TKCPKGGIA--AYST 1034
Db 752 AMGSSVYVNDIDRGRVGRV-----YLOGRPDARMPDDLSKWYVRNDKGMVFPNAPAT 805
QY 1035 ----SVNLTDGQVLSRFBWAYHKPLKNSODYTEALRAARELAANITADLRKVPCTDP 1088
Db 806 GKWEVGPSPLERYNGVPAMEILGEAPGLSSGD--AMAAVEE--IVKQLPKGVG-- 855
QY 1089 APEVFPYTTITNVEYQYLT-----ILPEGLFMLSIC-----LVPTFAYS 1127
Db 856 ----YSWTGYSYERLSSGQAPALYALSILVPLCLAALYESWSIPSVMVLVPLGT 909
QY 1128 CILLGLDRLSGLNLLSIVMILVDTVGFMAWDISYNVSLINL-----V 1172
Db 910 GALLATSMR-GLSNDVFFQVGLLTIGLSA-----KNAILIVEFAKELHEQCKGIVEAAI 963
QY 1173 SAVGMSVEFVSHITSFALSTPTWLERAKEATISMGSAVFACVAMTNLPGILVL-GLAK 1231
Db 964 EACRMRLRPVMTSLAFTILGVVP-----LAISTG-----AGSGSQHAIGTGVIGGMVT 1011
QY 1232 AOLIQIFPRL 1242
Db 1012 ATVLAIFWVPL 1022

RESULT 12

US-10-467-657-2648
; Sequence 2648, Application US/10467657

; Publication No. US20050260581A1

GENERAL INFORMATION:

; APPLICANT: CHIRON SpA

; APPLICANT: FONTANA Maria Rita

; APPLICANT: PIZZA Mariagrazia

; APPLICANT: MASIGNANI Vega

; APPLICANT: MONACI Elisabetta

; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/467,657

; CURRENT FILING DATE: 2003-08-11

; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12

; NUMBER OF SEQ ID NOS: 9218

; SOFTWARE: Seqwin99, version 1.04

; SEQ ID NO 2648

; LENGTH: 1067

; TYPE: PR

; ORGANISM: Neisseria gonorrhoeae

US-10-467-657-2648

Query Match

Best Local Similarity 19.4%; Score 134; DB 6; Length 1067;

Matches 185; Conservative 139; Mismatches 354; Indels 274; Gaps 44;

QY 452 RHLQW-SPEARQ--NISLQDICAPLNPDNTSLYDCCINSILQYPPQNRTLLLLLTANQT 508
Db 180 RAMRIVDPKLLQNTNLSPADVGS-LSAQNQISAGSIGSLPAV--RQGT---VTVATVT 233
QY 509 LMGQTSQVDKDHFLYCANAPLTFKQGTALALSCMADYCAPVFPFLAIGYKGYSEAE 568
Db 234 AGQQLGTAEFGNVILRANT-----DGSNIYLDKVAKVGLM-----EDSSST 277
QY 569 ALI-----MTPSLNYPAGDPRLAQAQKLEBAPLEBEMRAFORRMAGMFQVTPTAERSLE 622
Db 278 RLNGVNTTGMAMVLSN--SGN-AMATAKAVKERLAVLEKYFPQGMG--WKTPYDTSKFVE 332
QY 623 DEINRTAEDLPIFATSYIVIFLYISLALGSSYSSSRVWVDSKATLGLGVAVVLGAVMA 682
Db 333 ISIEKVIHTLIBAMVLVFMVYLPLONI--RYTLPTIIVVPISL---LGGFA----- 379
QY 683 AMGPFYSILGIRSSVLQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHVHIGRALGR 742
Db 380 ---FISYGM-SINVLTFMAMILVIGIVDDAIVVENVERIMAGEGLPPKEATKKAMQ 435
QY 743 VAPSM-----LCSLSEAICFFLGALTPMPAVTTPALTSLGLAVLDFLLQMSAFVALLSL 797
Db 436 ISGAVIGITAVLISVFPVPLAMPSGAAGNI--YKQFALTWASSI-----AFSAFLA--- 484
QY 798 DSKRQASRLDVCCCKQELPPPGQGGEL--LLGFFQKAYAPF-----LLHWIT 845
Db 485 ----TLTPALCATMLKTIPIKGHBEHKGFFGFWNKKPSTWTHGYGRVAKVLKTF 536
QY 846 RGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYLFLNRYFVGAPVFPVTT 905
Db 537 RMVYVIGLVVGV-----FLFMR-----LPTSFLPT 563
QY 906 LGYNFSBAGMNAICSSAGC-----NNFSPTQKIQYATEPPE--QSYLAIPASSWVDD----- 956
Db 564 EDQGFV---MVSQVLPAGATKERTDATLAQVTLAKSIPEIENIITVSGFSGSQNM 619
QY 957 ----FDW--LTPSSCCRLYIS-----GPNKDFPCPSTVN-----S 986
Db 620 AMGFALIKDMWERTASGSDAVAVAGKLTGMMGTLLKDGFGISVWPPPILELNGSGLSIN 679
QY 987 LNCLKNCMSITMGSVRPSVEQPHKYLPMF-----LNDPNIKCPKGGIAAYSTSVN 1037
Db 680 LQDRNNTGHTALLAKRNELIQMRASGLFPDPSVTVRAGGLEDSFQKIDINRAAAQAQGIS 739
QY 1038 LTSQGVLASRP--MAYHKPLKNSODYTEALRAARELAANITADLRK-----VP-- 1084
Db 740 FADIRTAASALSSSYVDFPNQGRQVRWVQADESDARMQPADILNLTVPKNSGVAVPLS 799
QY 1085 ----GTD-----PAFEVFPYTTITNVEYQYL----- 1106
Db 800 TIATVSWENGTEQSVRFNGYPSMKLSASPATGVTGQMAAAVQKMWDELGGGYSPFWGQ 859
QY 1107 ----TILPEGLFMLSICLV-----PTFAVSCLLAGL-----DLR 1136
Db 860 SREEAKGSGQTLILYGLAVAAVFLVLAALYESWSIPLAVILVPLGLIGAAAGVTGRNLF 919
QY 1137 SGLL-NLLSIVMILVDTVGFMAWDIS--YNAVSLINLV-----SAVMSVE-----F 1181
Db 920 EGLGVSFPSPANDIYFQVGFVVMGSLAKNAILIIEPAKDLQAQCKSAVEALEARLRF 979
QY 1182 VSHITRSPA--ISTKPTWLERAKEATI--SMGSAVFAGVAMTNLPGILVLGL 1229
Db 980 RPIIMTSPAFILGVVPLVYIAAGASSASQRAIGTTVFWGMLVGTLLSVFLVPL 1031

```

1121 -----VPTFAVSCLLGLDLRSGLLNLLSIWMLVDTVGFMAWLDISYNAVSLNL----- 1171
899 SVMLVPLGVVGALLAA--TFRGLTNDVYFQVGLTTIGLSA-----KNALLIVFPAKDL 951
1172 -----VSAGVMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMT 1219
952 MEKEGKGLIEATLEAVRMRLRPILMTSLAFILGWP-----LVISSG-----AGSGAQ 999
1220 NLPGLVL-GLAKAQLIQIFF 1239
1000 NAVGTGVMGGMVATILAIFF 1020

RESULT 14
US-11-242-459-16
; Sequence 16, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 16
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-16

Query Match 1.8%; Score 126.5; DB 7; Length 200;
Best Local Similarity 25.7%; Pred. No. 0.015; Indels 23; Gaps 8;
Matches 53; Conservative 38; Mismatches 92;

QY 612 QVTFATRSLSDEINRTTAED--LPFATSVIVIFLYISLALGYSYSSRVVMVDSKATILG 669
DB 3 QVLYGGTDLFYEVVRRTFNNDMLLAFISSCIAALVYILTSCSVFLSPFGI-----ASIG 57

QY 670 LGG-VAVVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVSVGADNIFIPVLEYQRLPR-R 727
DB 58 LSLVALFLXHVFGI---QYLGILNG-----VAAFVIVGIGVDDVFFVINTYQATHLE 109

QY 728 PQEPREVIHGRALGRVAPSMLLCSLSEALCPFLGALTMPAVRTFALTSGLAVI---LDF 784
DB 110 DPQLRMHTVQTAGK---ATPFTSLTTAAAYANVFSQIPAVHDFGLPMSLIVSCCLAV 166

QY 785 LQMSAFVALLSLDSKROEASRLDVC 810
DB 167 LVTPAALGLNSLYLAPLESSQTSQ 192

RESULT 15
US-10-793-626-2020
; Sequence 2020, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472

US-10-392-234A-14
; Sequence 14, Application US/10392234A
; Publication No. US20050255538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxseer, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acrAB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-392-234A-14

Query Match 1.8%; Score 127; DB 6; Length 1048;
Best Local Similarity 18.4%; Pred. No. 0.15;
Matches 147; Conservative 107; Mismatches 243; Indels 304; Gaps 34;

QY 624 EINRTAEDLPATSYIVIFLYISLALGYSYSSRVVMVDSKATL--GLGGVAVVLG--A 679
DB 339 EVVKTLVE-----AILVFLVMYLFQNF-----RATLIPTIAVEVLLGTFA 381

QY 680 VMAAMGFYSYLGIRSSLVILQVVPFLVLSVG--ADNIFVLEYQRYLPRRPGEPREVIH 737
DB 382 VLAAGFG-----SINTLTWFG-VMLAIGLLVDDAIVVENVERVMAEGLPPEATR 432

QY 738 RALGR-----VAPSMLLCSLSEALCPFLGALTMPAVRTFALTSGLAVILDFLLQMSAFV 792
DB 433 KSMGQIQGALVGIAMVLSAVFIPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALILTP 490

QY 793 ALLSLDSKROEASRLDVCCKVQKQELPPGQEGELLLGFFQKAYAPFLHWT----- 844
DB 491 ALCA-----TMLKPIQKSGHATTG--FFGFNFRMFDKSTHYTSDVGNILR 535

QY 845 TRGVVILLFLALF-----GVSLYSMCHISVGLDQE----- 874
DB 536 STGRYLVLYLIIVVGMAWLFVRLPSSFLPDEQGVFL-SMAQLPAGATQERTQKVDLWNT 594

QY 875 ---LALPKDS---YLLDVFLFLNR----- 892
DB 595 NYVLTREKDNVESVFAVNGFGFAGRGONTGIAFVSLKDWMSQRGPEENKVEAITARAMGYF 654

QY 893 -----YFEGVAPVYFTTLGYNFSSEAGNNAICSSAGCNFPSTQ----- 932
DB 655 SQIKDAMVFAFNLPAIVELGTATGDFDE-----LIDQAGLHEKLTQARNQLFGWVAQ 707

QY 933 -----KIQVATEPPEQSYLAIPAS-----SNVDDFDLWLTTP 963
DB 708 HPDVLTVRNGLEDTPQFKIDIDQEKALQALGVISIDINTTLGAANGSVNDFID---R 764

QY 964 SSCRLYISGNPKDKFCPTVNSLNCILKNCWITMGSRVPSVQFHKYLPWFNLDRPNLK 1023
DB 765 GRVKVYIMSEAKYRMLPEDIGK-----WYVRGSDGQM 797

QY 1024 CPKGGLAAYST-----SVNLTSQGVLSRFAWYHKPLKNSQDYTEALFAARELANIT 1077
DB 798 VP---PSAFSTSRWEYSGPLRYNGLPSLEILGQAAPGKST---GERMALMEELAKLP 851

QY 1078 ADLRKVPGTDPAFEPFYTITNVEYEQYLT-----ILPGLFMLSLCL----- 1120
DB 852 SGI-----CYDTGMSYQERLSGNQAPALYALISLVVFLCLAALYESWSIFP 898
```


THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:22:20 ; Search time 569 Seconds
(without alignments)
3235.086 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAI SNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:

- 1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pep:**
- 2: /cgn2_6/ptodata/1/paa/US066 COMB.pep:**
- 3: /cgn2_6/ptodata/1/paa/US073 COMB.pep:**
- 4: /cgn2_6/ptodata/1/paa/US074 COMB.pep:**
- 5: /cgn2_6/ptodata/1/paa/US075 COMB.pep:**
- 6: /cgn2_6/ptodata/1/paa/US076 COMB.pep:**
- 7: /cgn2_6/ptodata/1/paa/US077 COMB.pep:**
- 8: /cgn2_6/ptodata/1/paa/US078 COMB.pep:**
- 9: /cgn2_6/ptodata/1/paa/US079 COMB.pep:**
- 10: /cgn2_6/ptodata/1/paa/US080 COMB.pep:**
- 11: /cgn2_6/ptodata/1/paa/US081 COMB.pep:**
- 12: /cgn2_6/ptodata/1/paa/US082 COMB.pep:**
- 13: /cgn2_6/ptodata/1/paa/US083 COMB.pep:**
- 14: /cgn2_6/ptodata/1/paa/US084 COMB.pep:**
- 15: /cgn2_6/ptodata/1/paa/US085 COMB.pep:**
- 16: /cgn2_6/ptodata/1/paa/US086 COMB.pep:**
- 17: /cgn2_6/ptodata/1/paa/US087 COMB.pep:**
- 18: /cgn2_6/ptodata/1/paa/US088 COMB.pep:**
- 19: /cgn2_6/ptodata/1/paa/US089 COMB.pep:**
- 20: /cgn2_6/ptodata/1/paa/US090 COMB.pep:**
- 21: /cgn2_6/ptodata/1/paa/US091 COMB.pep:**
- 22: /cgn2_6/ptodata/1/paa/US092 COMB.pep:**
- 23: /cgn2_6/ptodata/1/paa/US093 COMB.pep:**
- 24: /cgn2_6/ptodata/1/paa/US094 COMB.pep:**
- 25: /cgn2_6/ptodata/1/paa/US095 COMB.pep:**
- 26: /cgn2_6/ptodata/1/paa/US096 COMB.pep:**
- 27: /cgn2_6/ptodata/1/paa/US097 COMB.pep:**
- 28: /cgn2_6/ptodata/1/paa/US098 COMB.pep:**
- 29: /cgn2_6/ptodata/1/paa/US099 COMB.pep:**
- 30: /cgn2_6/ptodata/1/paa/US100 COMB.pep:**
- 31: /cgn2_6/ptodata/1/paa/US101 COMB.pep:**
- 32: /cgn2_6/ptodata/1/paa/US102 COMB.pep:**
- 33: /cgn2_6/ptodata/1/paa/US103 COMB.pep:**
- 34: /cgn2_6/ptodata/1/paa/US104 COMB.pep:**
- 35: /cgn2_6/ptodata/1/paa/US105 COMB.pep:**
- 36: /cgn2_6/ptodata/1/paa/US106 COMB.pep:**
- 37: /cgn2_6/ptodata/1/paa/US107 COMB.pep:**
- 38: /cgn2_6/ptodata/1/paa/US108 COMB.pep:**
- 39: /cgn2_6/ptodata/1/paa/US109 COMB.pep:**
- 40: /cgn2_6/ptodata/1/paa/US110 COMB.pep:**
- 41: /cgn2_6/ptodata/1/paa/US111 COMB.pep:**
- 42: /cgn2_6/ptodata/1/paa/US112 COMB.pep:**
- 43: /cgn2_6/ptodata/1/paa/US114 COMB.pep:**

- 44: /cgn2_6/ptodata/1/paa/US600 COMB.pep:**
- 45: /cgn2_6/ptodata/1/paa/US601 COMB.pep:**
- 46: /cgn2_6/ptodata/1/paa/US602 COMB.pep:**
- 47: /cgn2_6/ptodata/1/paa/US603 COMB.pep:**
- 48: /cgn2_6/ptodata/1/paa/US604 COMB.pep:**
- 49: /cgn2_6/ptodata/1/paa/US605 COMB.pep:**
- 50: /cgn2_6/ptodata/1/paa/US606 COMB.pep:**
- 51: /cgn2_6/ptodata/1/paa/US607 COMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	1	PCT-US03-22467-4
2	6909	100.0	1332	1	PCT-US03-40113-4
3	6909	100.0	1332	1	PCT-US05-01469-4
4	6909	100.0	1332	36	US-10-621-758A-4
5	6909	100.0	1332	36	US-10-646-301A-4
6	6909	100.0	1332	36	US-10-663-208A-4
7	6909	100.0	1332	37	US-10-736-769-4
8	6909	100.0	1332	37	US-10-750-386-4
9	6896	99.8	1332	1	PCT-US01-04098A-1831
10	6896	99.8	1332	32	US-10-239-316-9
11	6896	99.8	1332	32	US-10-258-899A-1831
12	6896	99.8	1332	32	US-10-293-244-1831
13	6896	99.8	1332	42	US-11-218-141-1831
14	6896	99.8	1332	42	US-11-242-459-9
15	6872.5	99.5	1359	1	PCT-US01-04098A-1830
16	6872.5	99.5	1359	1	PCT-US03-22467-4
17	6872.5	99.5	1359	1	PCT-US03-40113-4
18	6872.5	99.5	1359	1	PCT-US05-01469-4
19	6872.5	99.5	1359	1	PCT-US05-27579-2
20	6872.5	99.5	1359	32	US-10-258-899A-1830
21	6872.5	99.5	1359	32	US-10-293-244-1830
22	6872.5	99.5	1359	36	US-10-621-758A-4
23	6872.5	99.5	1359	36	US-10-646-301A-4
24	6872.5	99.5	1359	36	US-10-663-208A-4
25	6872.5	99.5	1359	37	US-10-736-769-4
26	6872.5	99.5	1359	37	US-10-750-386-4
27	6872.5	99.5	1359	42	US-11-218-141-1830
28	6868.5	99.4	1359	31	US-10-170-205B-21686
29	6868.5	99.4	1359	48	US-60-453-050-8956
30	6868.5	99.4	1359	48	US-60-453-135-8956
31	6868.5	99.4	1359	48	US-60-466-412-8956
32	6536	94.6	1344	1	PCT-US01-08631-53052
33	6536	94.6	1344	34	US-10-450-763-53052
34	5421.5	78.5	1331	1	PCT-US03-22467-2
35	5421.5	78.5	1331	1	PCT-US03-40113-2
36	5421.5	78.5	1331	1	PCT-US05-01469-2
37	5421.5	78.5	1331	36	US-10-621-758A-2
38	5421.5	78.5	1331	36	US-10-646-301A-2
39	5421.5	78.5	1331	36	US-10-663-208A-2
40	5421.5	78.5	1331	37	US-10-736-769-2
41	5421.5	78.5	1331	37	US-10-750-386-2
42	5407	78.3	1333	1	PCT-US03-22467-12
43	5407	78.3	1333	1	PCT-US03-40113-12
44	5407	78.3	1333	1	PCT-US05-01469-12
45	5407	78.3	1333	1	PCT-US05-27579-3

ALIGNMENTS

RESULT 1
PCT-US03-22467-4
; Sequence 4, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation


```
QY 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFABQSDYSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFABQSDYSCSRVRPAA 180
QY 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
Db 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
QY 241 ARCNSQGDVATCSCQDCAASCAPAIARPAQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGDVATCSCQDCAASCAPAIARPAQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHDQHPFFRTNVTIAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHDQHPFFRTNVTIAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTFKOGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTFKOGTALAL 540
QY 541 SCHADYGAVFPFPLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLEM 600
Db 541 SCHADYGAVFPFPLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLEM 600
QY 601 RAFORMAGMFWOTFTAERSLDEINRTTAEDLPFATSVIYVIFLYISALGYSWSRV 660
Db 601 RAFORMAGMFWOTFTAERSLDEINRTTAEDLPFATSVIYVIFLYISALGYSWSRV 660
QY 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVADNIFIFVLE 720
QY 721 YQRLPRRPEPREVHIGRALGRVAPSMLLCSISEAICFFLGAITPMPAVRTTALTSGLA 780
Db 721 YQRLPRRPEPREVHIGRALGRVAPSMLLCSISEAICFFLGAITPMPAVRTTALTSGLA 780
QY 781 ILDELLOMSAFVALLSLDSKREASRLDYCCVKPQELPPPGQEGLLGFFQKAYAPFL 840
Db 781 ILDELLOMSAFVALLSLDSKREASRLDYCCVKPQELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNGNFQTKIQVATEFPPOSYLAIPASSWDDDFIDW 960
Db 901 YFVTTLGYNFSSEAGNNAICSSAGCNGNFQTKIQVATEFPPOSYLAIPASSWDDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKCPSVNSLNCIKCMSITMGSVRPSVEQFKHYPFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKCPSVNSLNCIKCMSITMGSVRPSVEQFKHYPFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAVHKPLKNSQDVTTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAVHKPLKNSQDVTTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAPFPPYPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLL 1140
Db 1081 RKVPGTDPAPFPPYPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLL 1140
QY 1141 NLLSIWMLVDTGFMALWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIWMLVDTGFMALWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMSGSAFVAGVAMTNLPGILVLGLAKAQLIQIIPFFRLNLLITLLGLLHGLVFLPV 1260
```

```
Db 1201 AKEATISMSGSAFVAGVAMTNLPGILVLGLAKAQLIQIIPFFRLNLLITLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDVNPALALEQKREAAVAVMVASCPNHPSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDVNPALALEQKREAAVAVMVASCPNHPSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
RESULT 3
PCT-US05-01469-4
; Sequence 4, Application PC/TUS0501469
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF
; FILE REFERENCE: 36134-PCT 074669.0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; PRIOR FILING DATE: 2005-01-25
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01469-4
Query Match 100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVLSLEASLSITKALLTRCPACDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVLSLEASLSITKALLTRCPACDNF 120
QY 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFABQSDYSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFABQSDYSCSRVRPAA 180
QY 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
Db 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
QY 241 ARCNSQGDVATCSCQDCAASCAPAIARPAQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGDVATCSCQDCAASCAPAIARPAQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHDQHPFFRTNVTIAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHDQHPFFRTNVTIAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTFKOGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTFKOGTALAL 540
QY 541 SCHADYGAVFPFPLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLEM 600
```

541 SCWADYGAVPFPFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKWEAEFLBEM 600
601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISIALGYSYSSWRV 660
601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISIALGYSYSSWRV 660
661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
721 YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTFALTSLAV 780
721 YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTFALTSLAV 780
781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLAIPASSVWDDFIDW 960
901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLAIPASSVWDDFIDW 960
961 LTPSSCCRLYISGPNKDFPCSTVNSLNCNKMSITMGSVRPSVEQFHKYLFWFLNDRP 1020
961 LTPSSCCRLYISGPNKDFPCSTVNSLNCNKMSITMGSVRPSVEQFHKYLFWFLNDRP 1020
1021 NTKCPKGGAAVSTSNLTSQVLAERFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
1021 NTKCPKGGAAVSTSNLTSQVLAERFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
1081 RKVPGTDPAFEPFYTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLSGILL 1140
1081 RKVPGTDPAFEPFYTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLSGILL 1140
1141 NLLSIVMILVDVTFGMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKTWLER 1200
1141 NLLSIVMILVDVTFGMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKTWLER 1200
1201 AKEATISGMSAVPAGVAMTNLPGILVLGLAKAQLIQIPFRNLNLTITLGLLHGLVFLPV 1260
1201 AKEATISGMSAVPAGVAMTNLPGILVLGLAKAQLIQIPFRNLNLTITLGLLHGLVFLPV 1260
1261 ILSYVGPDPVPALALEQKRAEBAVAVMVVASCNPHPSRVSTADNITYVNHSPFGSIKGAGA 1320
1261 ILSYVGPDPVPALALEQKRAEBAVAVMVVASCNPHPSRVSTADNITYVNHSPFGSIKGAGA 1320
1321 ISNFLPNNGROF 1332
1321 ISNFLPNNGROF 1332

RESULT 4

US-10-621-758A-4
; Sequence 4, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4
Query Match 100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGSLMTLSNVSCUSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGSLMTLSNVSCUSN 60
QY 61 TPARKITGDHILLLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHILLLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAGVTGCGYGSALCNAQRWLNPOGDTGNGLAPLDITFHLLERQAVGSGIQLNEGV 240
DB 181 ATLAGVTGCGYGSALCNAQRWLNPOGDTGNGLAPLDITFHLLERQAVGSGIQLNEGV 240
QY 241 ARCHESQDDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIIILCSVFAVVTI 300
DB 241 ARCHESQDDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
DB 301 LLVGFRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPVELVSNAPNSQARSEKAFHDOHGPFRTNQVILITAPNRSY 420
DB 361 IPVVALAAGLVFTELTTPVELVSNAPNSQARSEKAFHDOHGPFRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLLELELQERLRLHLOWSPSEARQNISLODICVAPLNPDNT 480
DB 421 RYDLSLLGPKNPSGILDLLELELQERLRLHLOWSPSEARQNISLODICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVMDKDHPLYCANAPLTPKDGDTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVMDKDHPLYCANAPLTPKDGDTALAL 540
QY 541 SCWADYGAVPFPFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKWEAEFLBEM 600
DB 541 SCWADYGAVPFPFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKWEAEFLBEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISIALGYSYSSWRV 660
DB 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISIALGYSYSSWRV 660
QY 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
DB 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTFALTSLAV 780
DB 721 YQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTFALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
QY 841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLAIPASSVWDDFIDW 960
DB 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLAIPASSVWDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFPCSTVNSLNCNKMSITMGSVRPSVEQFHKYLFWFLNDRP 1020

```
Db 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKNCMSITMGSVRPSVQFHKYLPWFNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAFEVPPYITITNVFYEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVPPYITITNVFYEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIWMLVDTVGFMAWDISYNAVSLNLSAVGMSVEFVSHITRSPAIKPTWLER 1200
Db 1141 NLLSIWMLVDTVGFMAWDISYNAVSLNLSAVGMSVEFVSHITRSPAIKPTWLER 1200
Qy 1201 AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRNLMLITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRNLMLITLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPFGSIKGAGA 1320
Qy 1321 ISNFPNNGRQF 1332
Db 1321 ISNFPNNGRQF 1332

RESULT 5
US-10-646-301A-4
; Sequence 4, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-4

Query Match 100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPGLSGMLTSLNSVCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPGLSGMLTSLNSVCLSN 60
Qy 61 TPARKITGDHLLILLOKICPRLVTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLVTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNOSLFINVTRVAQIGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
Db 121 VNLHCHTCSNOSLFINVTRVAQIGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
Qy 181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCNSQGDVDATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
```

```
Db 241 ARCNSQGDVDATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGFVAPARDKSMVDQKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSMVDQKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGFPFRFTNOVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGFPFRFTNOVILTAPNRSSY 420
Qy 421 RYDLSLLGPKNSGILDLIDLLELLELQRLHRLQWSPEAQRNISLODICVAPLNPDMT 480
Db 421 RYDLSLLGPKNSGILDLIDLLELLELQRLHRLQWSPEAQRNISLODICVAPLNPDMT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Qy 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLMEAEFLSEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLMEAEFLSEM 600
Qy 601 RAFORMMAGMFOVITFAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALGYSWSRV 660
Db 601 RAFORMMAGMFOVITFAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALGYSWSRV 660
Qy 661 MYDSKATILGLGVAVVILGAVMAAGFPFSLGRSSIVILQVVPFLVLSVGAADNIFIVLE 720
Db 661 MYDSKATILGLGVAVVILGAVMAAGFPFSLGRSSIVILQVVPFLVLSVGAADNIFIVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEALICFFELGALTMPAVRTALTSLGLAV 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEALICFFELGALTMPAVRTALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQOELPPPGQEGELLGFGQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQOELPPPGQEGELLGFGQKAYAPFL 840
Qy 841 LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTLYGNFSSEAGMNAICSSAGCNPFSTQKIQVATEFPFQSYLAIPASSWVDDFDW 960
Db 901 YFVTTLYGNFSSEAGMNAICSSAGCNPFSTQKIQVATEFPFQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKNCMSITMGSVRPSVQFHKYLPWFNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKNCMSITMGSVRPSVQFHKYLPWFNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAFEVPPYITITNVFYEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVPPYITITNVFYEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIWMLVDTVGFMAWDISYNAVSLNLSAVGMSVEFVSHITRSPAIKPTWLER 1200
Db 1141 NLLSIWMLVDTVGFMAWDISYNAVSLNLSAVGMSVEFVSHITRSPAIKPTWLER 1200
Qy 1201 AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRNLMLITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRNLMLITLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPFGSIKGAGA 1320
Qy 1321 ISNFPNNGRQF 1332
Db 1321 ISNFPNNGRQF 1332
```

```
Db      1321 ISNFLPNNGRQF 1332

RESULT 6
US-10-663-208A-4
; Sequence 4, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; PRIOR FILING DATE: 2003-09-16
; PRIOR FILING DATE: 2003-07-19
; PRIOR FILING DATE: 2002-07-19
; PRIOR FILING DATE: 2003-07-17
; PRIOR FILING DATE: 2003-07-17
; PRIOR FILING DATE: 2003-07-17
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-208A-4

Query Match      100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFVDECGKNPELSGSLMTLSNVCSLSN 60
Db      1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFVDECGKNPELSGSLMTLSNVCSLSN 60

Qy      61 TPARKITGHLILLQKICPRLYTGNPTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
Db      61 TPARKITGHLILLQKICPRLYTGNPTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120

Qy      121 VNLHCHNTCSNQSLFINTRVAQLGAGOLPAVVAEAFYQHSFAEQSDSCSRVRPAA 180
Db      121 VNLHCHNTCSNQSLFINTRVAQLGAGOLPAVVAEAFYQHSFAEQSDSCSRVRPAA 180

Qy      181 ATLAVGTMCVGYGSALCNARQWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db      181 ATLAVGTMCVGYGSALCNARQWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240

Qy      241 ARCNEQGGDDVATCSCQDCAACPATARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db      241 ARCNEQGGDDVATCSCQDCAACPATARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

Qy      301 LLVGRVAPARDKSKMVDPKGTSLSDKLSFSHTLLGQFFQGWGTWASWPLTILVLSV 360
Db      301 LLVGRVAPARDKSKMVDPKGTSLSDKLSFSHTLLGQFFQGWGTWASWPLTILVLSV 360

Qy      361 IPVVALAAGLVFTLTTPVELMSAPNSQARSEKAFHDQHFQFFRTNVTAPNRSY 420
Db      361 IPVVALAAGLVFTLTTPVELMSAPNSQARSEKAFHDQHFQFFRTNVTAPNRSY 420

Qy      421 RYDSSLGPNKPNFGSILDLILLLELLELRLHQLVMSPEAQNRISLQDICYAPLNPNT 480
Db      421 RYDSSLGPNKPNFGSILDLILLLELLELRLHQLVMSPEAQNRISLQDICYAPLNPNT 480

Qy      481 SLVDDCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGITALAL 540
Db      481 SLVDDCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGITALAL 540

Qy      541 SCMDYGPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db      541 SCMDYGPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600

Qy      601 RAFORMMAGMQVFTTFAERSLEDEINRTTAEDLPFATSYIVIVFLYISLALGSYSSMSRV 660
Db      601 RAFORMMAGMQVFTTFAERSLEDEINRTTAEDLPFATSYIVIVFLYISLALGSYSSMSRV 660

Qy      661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGRSSVLQVVPFLVLSVGDADNIFIFVLE 720
Db      661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGRSSVLQVVPFLVLSVGDADNIFIFVLE 720

Qy      721 YQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLIGALTPMPAVRTFALTSLGLAV 780
Db      721 YQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLIGALTPMPAVRTFALTSLGLAV 780

Qy      781 ILDFELLQMSAFVALLSLDSKQEAASRLDVCCKVPQELPPPGQEGGLLGFQKAYAPFL 840
Db      781 ILDFELLQMSAFVALLSLDSKQEAASRLDVCCKVPQELPPPGQEGGLLGFQKAYAPFL 840

Qy      841 LHWITRGVWVLLFLALFGVSLYSCHISVGLDQBLALPKDSYLLDYFLFLNRYFVGAPV 900
Db      841 LHWITRGVWVLLFLALFGVSLYSCHISVGLDQBLALPKDSYLLDYFLFLNRYFVGAPV 900

Qy      901 YFVTTGLGYNFSSAGMNAICSSAGCNPFSTQKIYATEFPEQSYLAIPASSWVDDFIDW 960
Db      901 YFVTTGLGYNFSSAGMNAICSSAGCNPFSTQKIYATEFPEQSYLAIPASSWVDDFIDW 960

Qy      961 LTPSSCCRLYISGPNKOKFCPSTVNSLNCNKMCSITMGSVRPSVEQPHKYLWFLNDRP 1020
Db      961 LTPSSCCRLYISGPNKOKFCPSTVNSLNCNKMCSITMGSVRPSVEQPHKYLWFLNDRP 1020

Qy      1021 NIKCPKGLAAYSTVNLTSQGVLAHSPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db      1021 NIKCPKGLAAYSTVNLTSQGVLAHSPMAYHKPLKNSQDYTEALRAARELAANITADL 1080

Qy      1081 RKVPGTDPAFEPVPTTINVFVEQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db      1081 RKVPGTDPAFEPVPTTINVFVEQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140

Qy      1141 NLLSIVMLVDTVGPMAWLDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200
Db      1141 NLLSIVMLVDTVGPMAWLDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200

Qy      1201 AKEATISMGSAVFAVAMTNLPGIILVGLAKAQLIQIFFRLNLLITLGLLHGLVFLPV 1260
Db      1201 AKEATISMGSAVFAVAMTNLPGIILVGLAKAQLIQIFFRLNLLITLGLLHGLVFLPV 1260

Qy      1261 ILSYVGPVDPNPALEOKRAEAAVAVMVASCPNHPSPRVSTADNIVNHSPEGSIKGAGA 1320
Db      1261 ILSYVGPVDPNPALEOKRAEAAVAVMVASCPNHPSPRVSTADNIVNHSPEGSIKGAGA 1320

Qy      1321 ISNFLPNNGRQF 1332
Db      1321 ISNFLPNNGRQF 1332

RESULT 7
US-10-736-769-4
; Sequence 4, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
```



```

; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-4

Query Match      100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGHWLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGHWLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICPRLTYGNTQACCSAKQLVSEASLSITKALLTRCPACSNF 120
DB 61 TPARKITGDHLLILLOKICPRLTYGNTQACCSAKQLVSEASLSITKALLTRCPACSNF 120
QY 121 VNLHCHNTCSNQSIFINVTVAQAGOLPAVVAEAFYQHSFAEQSDSCSRVRVPA 180
DB 121 VNLHCHNTCSNQSIFINVTVAQAGOLPAVVAEAFYQHSFAEQSDSCSRVRVPA 180
QY 181 ATLVGTMCGVYGSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATLVGTMCGVYGSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQDDVATCSCDCAAPARQALDSTFYLGOMPGSLVLIILCSVAVVTI 300
DB 241 ARCNESQDDVATCSCDCAAPARQALDSTFYLGOMPGSLVLIILCSVAVVTI 300
QY 301 LLVGRVAPARDKRWDPKGTSTLSKLSFSTHTLLGQFQCGWTWASWPLTILVLSV 360
DB 301 LLVGRVAPARDKRWDPKGTSTLSKLSFSTHTLLGQFQCGWTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFPHDQHPFPRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFPHDQHPFPRTNQVILTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELELQERLHLQVNSPEAQRNLSIODICYAPLNPNT 480
DB 421 RYDSLILGPKNPSGILDLLELELQERLHLQVNSPEAQRNLSIODICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLPCANAPLTKDGTAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLPCANAPLTKDGTAL 540
QY 541 SCMADYGAPVFPFLAIGGKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEAPLEEM 600
DB 541 SCMADYGAPVFPFLAIGGKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEAPLEEM 600
QY 601 RAPORMAGMFOVTTAERSLEDEINRTAEDLPATSYIVIFLYISLAGSYSSWSRV 660
DB 601 RAPORMAGMFOVTTAERSLEDEINRTAEDLPATSYIVIFLYISLAGSYSSWSRV 660
QY 661 MYDSKATLGLGVAVVLGAVMAAMGPFVSLGRSSILVILQVVPVLVSGADNIFVLE 720
DB 661 MYDSKATLGLGVAVVLGAVMAAMGPFVSLGRSSILVILQVVPVLVSGADNIFVLE 720
QY 721 YQRLPRRPGEPREHIGRALGRVAFSMLLCSLSEACFFLGLATPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREHIGRALGRVAFSMLLCSLSEACFFLGLATPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQWSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQSGELLGFPQKAYAPPL 840
DB 781 ILDFLLQWSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQSGELLGFPQKAYAPPL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
QY 901 YFVTTLTGYNFSSEAGNNAICSSAGCWNFSFTQKIQVATEFPQSYLAIPASSWVDDFDW 960
```

```

DB 901 YFVTTLTGYNFSSEAGNNAICSSAGCWNFSFTQKIQVATEFPQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNLCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNLCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
QY 1021 NIKPKGGLAAYSTSVNLTSDQVILASRFMAYHKPKLNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKPKGGLAAYSTSVNLTSDQVILASRFMAYHKPKLNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVPPYTIITNVFYEQYLITLPEGLFMLSICLVPTFAVSCLLGLDRLSGLL 1140
DB 1081 RKVPGTDPAPFVPPYTIITNVFYEQYLITLPEGLFMLSICLVPTFAVSCLLGLDRLSGLL 1140
QY 1141 NLLSVIMLIVDTVGFPMALWDISYNAVSLINLSVAGMSVEFYVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSVIMLIVDTVGFPMALWDISYNAVSLINLSVAGMSVEFYVSHITRSPAISTKPTWLER 1200
QY 1201 AKREATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFRLNLNLTLLGLLHGLVFLPV 1260
DB 1201 AKREATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFRLNLNLTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALEQKRABEAVAAVMVASCPNHPSPRVSTADNIYVNHSPGSIKGAGA 1320
DB 1261 ILSYVGPDPVNPALALEQKRABEAVAAVMVASCPNHPSPRVSTADNIYVNHSPGSIKGAGA 1320
QY 1321 ISNPLPNNNGROP 1332
DB 1321 ISNPLPNNNGROP 1332
```

RESULT 8

```

US-10-750-386-4
; Sequence 4, Application US/10750386
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; APPLICANT: Chapman, Kevin
; APPLICANT: Goulet, Mark
; APPLICANT: Ujjainwalla, Feroze
; APPLICANT: Altmann, Scott W
; APPLICANT: Davis, Chip
; APPLICANT: Bull, Herb
; APPLICANT: Thornberry, Nancy A
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONIST
; FILE REFERENCE: A36104 074669 0103
; CURRENT APPLICATION NUMBER: US/10/750,386
; CURRENT FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-750-386-4
```

```

Query Match      100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGHWLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGHWLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICPRLTYGNTQACCSAKQLVSEASLSITKALLTRCPACSNF 120
DB 61 TPARKITGDHLLILLOKICPRLTYGNTQACCSAKQLVSEASLSITKALLTRCPACSNF 120
QY 121 VNLHCHNTCSNQSIFINVTVAQAGOLPAVVAEAFYQHSFAEQSDSCSRVRVPA 180
DB 121 VNLHCHNTCSNQSIFINVTVAQAGOLPAVVAEAFYQHSFAEQSDSCSRVRVPA 180
QY 181 ATLVGTMCGVYGSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
```

Db	181	ATLAVGTGCGVGSALCNAQRWLNFGDITGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240
Qy	241	ARCNSQGDVATCSQDCAASCAPAIARPAQALDSTFYLQMGFGLVLIILCSVFAVVTI	300
Db	241	ARCNSQGDVATCSQDCAASCAPAIARPAQALDSTFYLQMGFGLVLIILCSVFAVVTI	300
Qy	301	LLVGRFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Db	301	LLVGRFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHPGFFRFTNQVILITAPNRSY	420
Db	361	IPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHPGFFRFTNQVILITAPNRSY	420
Qy	421	RYDSLILGPNKFSIGILDLLLELLELQBLRHLQVWSPQAQRTSLQDICYAPLNPDNT	480
Db	421	RYDSLILGPNKFSIGILDLLLELLELQBLRHLQVWSPQAQRTSLQDICYAPLNPDNT	480
Qy	481	SLYDCCINSILQYFQNNRTLLALLTANQTLMGOTSQVDWKDHFLYCANAPLTFKDGTAL	540
Db	481	SLYDCCINSILQYFQNNRTLLALLTANQTLMGOTSQVDWKDHFLYCANAPLTFKDGTAL	540
Qy	541	SCWADYGAPVFPFLAIGGKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWAEFLBEM	600
Db	541	SCWADYGAPVFPFLAIGGKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWAEFLBEM	600
Qy	601	RAFQRMAGMFQVTFTAERSLDEINRTTAEDLPIPATSIYIVFLYISALGSYSWSRV	660
Db	601	RAFQRMAGMFQVTFTAERSLDEINRTTAEDLPIPATSIYIVFLYISALGSYSWSRV	660
Qy	661	MVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVVPFLVLSVGADNIFVLE	720
Db	661	MVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVVPFLVLSVGADNIFVLE	720
Qy	721	YQRLPRPGEPREVHIGRALGRVAPSPMLLCSLSEALICFFELGALTMPAPVTRTALTSGLA	780
Db	721	YQRLPRPGEPREVHIGRALGRVAPSPMLLCSLSEALICFFELGALTMPAPVTRTALTSGLA	780
Qy	781	ILDFLQMSAFVALLSLDSKROEASRLDVCCCKVQOELPPPGQEGGLLIGFFQKAYAPPL	840
Db	781	ILDFLQMSAFVALLSLDSKROEASRLDVCCCKVQOELPPPGQEGGLLIGFFQKAYAPPL	840
Qy	841	LHWITRGVLLFLALFGVSLYSMCHI SVGLQDELALPKDSYLLDYFLPLNRYFEVGAPV	900
Db	841	LHWITRGVLLFLALFGVSLYSMCHI SVGLQDELALPKDSYLLDYFLPLNRYFEVGAPV	900
Qy	901	YFVTTILGYNFSSEAGMNAICSSAGCNMFSTQKIQYATEFPFQSYLAIPASSWVDVDFIDW	960
Db	901	YFVTTILGYNFSSEAGMNAICSSAGCNMFSTQKIQYATEFPFQSYLAIPASSWVDVDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDKFCPTVNSLNLCKNCSITMGSVRPSVEQHKYLPWFINDRP	1020
Db	961	LTPSSCCRLYISGPNKDKFCPTVNSLNLCKNCSITMGSVRPSVEQHKYLPWFINDRP	1020
Qy	1021	NIKCPKGLAAYSTSVNLSDGQVLASRFMAHYKPLKNSQDYTEALRAARELAANITADL	1080
Db	1021	NIKCPKGLAAYSTSVNLSDGQVLASRFMAHYKPLKNSQDYTEALRAARELAANITADL	1080
Qy	1081	RKVPDTPAPEFPVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGILL	1140
Db	1081	RKVPDTPAPEFPVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGILL	1140
Qy	1141	NLLSIWMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKTWLER	1200
Db	1141	NLLSIWMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKTWLER	1200
Qy	1201	AKEATISMGSAVPAFVAMTNLPGLVGLAKAQLIQIFFFRNLNLTITLGLLHGLVFLPV	1260
Db	1201	AKEATISMGSAVPAFVAMTNLPGLVGLAKAQLIQIFFFRNLNLTITLGLLHGLVFLPV	1260
Qy	1261	ILSYGPDVNPALALEQKRAEBAVAAMVASCNPHPRSRVSTADNTIYNVHSPFGSIKGAGA	1320

Db	1261	ILSYGPDVNPALALEQKRAEBAVAAMVASCNPHPRSRVSTADNTIYNVHSPFGSIKGAGA	1320
Qy	1321	ISNFLPNNGRQF	1332
Db	1321	ISNFLPNNGRQF	1332
RESULT 9			
PCT-US01-04098A-1831			
; Sequence 1831, Application PC/TUS0104098A			
; GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc.			
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides			
; FILE REFERENCE: 21272-029			
; CURRENT APPLICATION NUMBER: PCT/US01/04098A			
; CURRENT FILING DATE: 2001-02-05			
; PRIOR APPLICATION NUMBER: Not Yet Assigned			
; PRIOR FILING DATE: 2001-01-30			
; PRIOR APPLICATION NUMBER: 09/728,422			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: 09/693,325			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 09/663,561			
; PRIOR FILING DATE: 2000-09-15			
; PRIOR APPLICATION NUMBER: 09/654,936			
; PRIOR FILING DATE: 2000-09-01			
; PRIOR APPLICATION NUMBER: 09/620,325			
; PRIOR FILING DATE: 2000-07-19			
; PRIOR APPLICATION NUMBER: 09/598,075			
; PRIOR FILING DATE: 2000-06-20			
; PRIOR APPLICATION NUMBER: 09/560,875			
; PRIOR FILING DATE: 2000-04-27			
; PRIOR APPLICATION NUMBER: 09/496,914			
; PRIOR FILING DATE: 2000-02-03			
; NUMBER OF SEQ ID NOS: 3960			
; SOFTWARE: Custom			
; SEQ ID NO 1831			
; LENGTH: 1332			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
PCT-US01-04098A-1831			
Query Match 99.8%; Score 6896; DB 1; Length 1332;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN	60
Qy	61	TPARKITGDHLLILQKICPRLYTGNTQACCSAKOLVSEASLSTTKALLTRCPACSNF	120
Db	61	TPARKITGDHLLILQKICPRLYTGNTQACCSAKOLVSEASLSTTKALLTRCPACSNF	120
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA	180
Db	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA	180
Qy	181	ATLAVGTWCVGYGSALCNAQRWLNFPQGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240
Db	181	ATLAVGTWCVGYGSALCNAQRWLNFPQGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240
Qy	241	ARCHESQGDVATCSQDCAASCAPAIARPAQALDSTFYLQMGFGLVLIILCSVFAVVTI	300
Db	241	ARCHESQGDVATCSQDCAASCAPAIARPAQALDSTFYLQMGFGLVLIILCSVFAVVTI	300
Qy	301	LLVGRFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Db	301	LLVGRFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHPGFFRFTNQVILITAPNRSY	420
Db	361	IPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHPGFFRFTNQVILITAPNRSY	420

```
QY 421 RYDSSLGPNKPSGILDLDDLLLELLELQRLRLHQLWSPQAQRNLSLQDICYAPLNPNT 480
Db 421 RYDSSLGPNKPSGILDLDDLLLELLELQRLRLHQLWSPQAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTPKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSWSRV 660
Db 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSWSRV 660
QY 661 MYDSKATLGLGGVAVVILGAVMAAMGPFYSYLGRSSLVILQVVPFLVLSGADNIPFVLE 720
Db 661 MYDSKATLGLGGVAVVILGAVMAAMGPFYSYLGRSSLVILQVVPFLVLSGADNIPFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQEGGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQEGGLLGFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLQDELALPKDSYLLDYFLFNRYFVGAVP 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLQDELALPKDSYLLDYFLFNRYFVGAVP 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSTQKI QYATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSTQKI QYATEFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYIISGPNKDFCPSTVNSLNCNKCMSITMGSVRPSVEQFHKLIPWFLNDRP 1020
Db 961 LTPSSCCRLYIISGPNKDFCPSTVNSLNCNKCMSITMGSVRPSVEQFHKLIPWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVFPYITITNVFYEQYLTILPEGLFMLSCLVPTPEAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAPFVFPYITITNVFYEQYLTILPEGLFMLSCLVPTPEAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMILVDTVGPMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGPMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVPAGVAMTNLFGILVLGAKAQLQIIPFRLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMSGSAVPAGVAMTNLFGILVLGAKAQLQIIPFRLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALEQKRAEBAVAVMVASCPNHPSTVSTADNIVNHSFSGSIKGAGA 1320
Db 1261 ILSYVGPDPVNPALALEQKRAEBAVAVMVASCPNHPSTVSTADNIVNHSFSGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
```

RESULT 10
US-10-239-316-9
; Sequence 9, Application US/10239316
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama

; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
; US-10-239-316-9

Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MABAGLRGWLWALLRLAQSEPYTTIHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MABAGLRGWLWALLRLAQSEPYTTIHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICERLYTGNTQACCSAKQLVSLSEASISITKALLTRCPACSDNP 120
Db 61 TPARKITGDHLILLOKICERLYTGNTQACCSAKQLVSLSEASISITKALLTRCPACSDNP 120
QY 121 VNLHCHTCSNQSIFINVRVAQLGAGQLPAVAYEAFYQHSFABQSVDSSSRVRPAA 180
Db 121 VNLHCHTCSNQSIFINVRVAQLGAGQLPAVAYEAFYQHSFABQSVDSSSRVRPAA 180
QY 181 ATLVAGTMCVGVYSALCNAQRWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db 181 ATLVAGTMCVGVYSALCNAQRWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNEQGDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVPAVVTI 300
QY 301 LLAGFRVAPARDKSKWVDPKGTSLSDKLSPTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLAGFRVAPARDKSKWVDPKGTSLSDKLSPTHTLLGQFFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTPDVELWSPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTPDVELWSPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSSY 420
QY 421 RYDSSLGPNKPSGILDLDDLLLELLELQRLRLHQLWSPQAQRNLSLQDICYAPLNPNT 480
Db 421 RYDSSLGPNKPSGILDLDDLLLELLELQRLRLHQLWSPQAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTPKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSWSRV 660
Db 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSWSRV 660
QY 661 MYDSKATLGLGGVAVVILGAVMAAMGPFYSYLGRSSLVILQVVPFLVLSGADNIPFVLE 720
Db 661 MYDSKATLGLGGVAVVILGAVMAAMGPFYSYLGRSSLVILQVVPFLVLSGADNIPFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQEGGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQEGGLLGFQKAYAPFL 840

Db 781 ILDFLLQMSAFVALLSLDSKQESASRLDVCCVKPQELPPRQGBGLLGFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSVLAIPASSWDDDFIDW 960
Db 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSVLAIPASSWDDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNLCNMSITMGSRPSVEQFHKYLPFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNLCNMSITMGSRPSVEQFHKYLPFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVGTDPAFEPFPYTIITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVGTDPAFEPFPYTIITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSVMLIVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSVMLIVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGLILVLGLAKAQLIQIIPFFRLNLITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGLILVLGLAKAQLIQIIPFFRLNLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVDPNPALEOKBAEAAVAVMVASCPNHPSPRVSTADNLYVNHSPSGSKGAGA 1320
Db 1261 ILSYVGPVDPNPALEOKBAEAAVAVMVASCPNHPSPRVSTADNLYVNHSPSGSKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 11

US-10-258-899A-1831
; Sequence 1831. Application US/10258899A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Cao, Yicheng
; APPLICANT: Ma, Yunqueing
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Xue, Aidong
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Goodrich, Ryle
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 787CIP2-2G/US
; CURRENT APPLICATION NUMBER: US/10/258, 899A
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: PCT/US01/04098
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/774,434
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-899A-1831

Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGPTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGPTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNPQSLFINVTRVAQGLGQLPVAVAYEAFYOHSPAEQSYDSCSRVRVPA 180
Db 121 VNLCHNTCSNPQSLFINVTRVAQGLGQLPVAVAYEAFYOHSPAEQSYDSCSRVRVPA 180
QY 181 ATAVGTCMGVYGSALCNAQRLNFGQDGTGNGLAPLDITTFHLEPGQAVGSGIQPLNEGV 240
Db 181 ATAVGTCMGVYGSALCNAQRLNFGQDGTGNGLAPLDITTFHLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVVATCSCQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVVATCSCQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFCQGWCTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMWDPKGTSLSDKLSFSTHTLLGQFCQGWCTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420
QY 421 RYDSLGLGPKNPSGILDLDLLELELERLHLQVMSPEAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSLGLGPKNPSGILDLDLLELELERLHLQVMSPEAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKHFLYCANAPLTKDGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKHFLYCANAPLTKDGTALAL 540
QY 541 SCWADYGAPVFPFLLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFWQVTFMAERSLEDEINRTTAEDLPFATSYIIVFLYISALGSYSWSRV 660
Db 601 RAFQRMAGMFWQVTFMAERSLEDEINRTTAEDLPFATSYIIVFLYISALGSYSWSRV 660
QY 661 MYDSKATLGLGCVAVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
Db 661 MYDSKATLGLGCVAVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSGADNIFIFVLE 720

```
QY 721 YORLPBPGEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSLAV 780
Db 721 YORLPBPGEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPQGGELLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPQGGELLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNMFSTQKIQYATEFFPQSYLAIPASSWVDDPIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNMFSTQKIQYATEFFPQSYLAIPASSWVDDPIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNLCKNMSITMGSVPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNLCKNMSITMGSVPSVEQFHXYLPWFINDRP 1020
QY 1021 NIKCPKGLAAVSTSNLTSDQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAVSTSNLTSDQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAFEPFPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL 1140
Db 1081 RKVPGTDPAFEPFPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL 1140
QY 1141 NLLSTVMTLLVDVTFGMALWDISYNAVSLNLINLSAVGMSVEFVSHLTRSPAISTKPTWLER 1200
Db 1141 NLLSTVMTLLVDVTFGMALWDISYNAVSLNLINLSAVGMSVEFVSHLTRSPAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLTILGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLTILGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVPALALEQKAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNLPNNGRQF 1332
Db 1321 ISNLPNNGRQF 1332
```

RESULT 12

```
US-10-293-244-1831
; Sequence 1831, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
```

```
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1831
```

```
Query Match          99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MAEAGLRGWLWALLRLRLAQSPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLRLAQSPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLLOKICPRLVTGNTQACCSAKOLVLEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLLOKICPRLVTGNTQACCSAKOLVLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSIFINVRVAQLGAGOLPAVVAEYAFYQHSFABQSVDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSIFINVRVAQLGAGOLPAVVAEYAFYQHSFABQSVDSCSRVRPAA 180
QY 181 ATLVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFFHLEPGQAVGSGIQPLNEGV 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFFHLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILCSVPFVVTI 300
Db 241 ARCNEQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILCSVPFVVTI 300
QY 301 LLVGRVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFETLTPDVELWSAPNSQARSEKAFHDQHPGFPRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFETLTPDVELWSAPNSQARSEKAFHDQHPGFPRTNQVILTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLDDLELELELELELELELELELELELELELELELELELELE 480
Db 421 RYDSLILGPKNPSGILDLDDLELELELELELELELELELELELELELELELELELELE 480
QY 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDWKDHFYLCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDWKDHFYLCANAPLTFKDGFTALAL 540
QY 541 SCMDYGAVPFPFLAIGGYKGYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCMDYGAVPFPFLAIGGYKGYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAPQRMAGMFQVTPAERSLEDEINRTTAEADLPFATSYIVIFLYISALSGYSWSRV 660
Db 601 RAPQRMAGMFQVTPAERSLEDEINRTTAEADLPFATSYIVIFLYISALSGYSWSRV 660
QY 661 MVDKATIGLGGVAVVLGAVNMAAGFVSILGRSSLVILQVVPFLVLSVGAONIFIFVLE 720
Db 661 MVDKATIGLGGVAVVLGAVNMAAGFVSILGRSSLVILQVVPFLVLSVGAONIFIFVLE 720
QY 721 YQRLPRRPGEPREHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSLAV 780
Db 721 YQRLPRRPGEPREHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPQGGELLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPQGGELLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNMFSTQKIQYATEFFPQSYLAIPASSWVDDPIDW 960
```

Db 901 YFVITLGYNFSSEAGNNAICSSAGCNNSFTQKIQYATEFFEQSVLAIPASSWDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Qy 1021 NIKCPKGGGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGGGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIVMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRFSFAISTKPTWLER 1200
Db 1141 NLLSIVMLVDTVGFMAWGISYNAVSLINLSAVGMSVEFVSHITRFSFAISTKPTWLER 1200
Qy 1201 AKEATISGMSAVFAGVAMTNLPGLVLGLAKAQLIOIPFRLNALITLLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGLVLGLAKAQLIOIPFRLNALITLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPDPNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 13
US-11-218-141-1831
; Sequence 1831, Application US/11218141
; GENERAL INFORMATION:
; APPLICANT: Yang, Y. Tom et al
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/11/218,141
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-218-141-1831

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECCKNPSELGSLMTLSNVCLSN 60

Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECCKNPSELGSLMTLSNVCLSN 60
Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITTKALLTRCPACSNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITTKALLTRCPACSNF 120
Qy 121 VNLCHNTCSPNQSLFTNVTRVAOLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLCHNTCSPNQSLFTNVTRVAOLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAVGTCMGVYGALCNQRLNFPQDGTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
Db 181 ATLAVGTCMGVYGALCNQRLNFPQDGTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
Qy 241 ARCHESQDDVATCSCQDCAAPALPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCHESQDDVATCSCQDCAAPALPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSPSTHTLLGQFFQGMGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSPSTHTLLGQFFQGMGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGPPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGPPFRTNQVILTAPNRSY 420
Qy 421 RYDSLLLGPKNFGSILDLLELELERLRLHQVWSPQARNISLODICVAPLNPDNT 480
Db 421 RYDSLLLGPKNFGSILDLLELELERLRLHQVWSPQARNISLODICVAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGITAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGITAL 540
Qy 541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Qy 601 RAFQRMAGMQVTFPAERSLEDEINRTAEDLPFATSYIVIFLYISALGYSWSVRV 660
Db 601 RAFQRMAGMQVTFPAERSLEDEINRTAEDLPFATSYIVIFLYISALGYSWSVRV 660
Qy 661 MYDSKATILGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVADNIFIFVLE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFPLGALTMPAVRTALTSLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFPLGALTMPAVRTALTSLAV 780
Qy 781 ILDFLLQMSAFVALLSLSKROEASRLDVCCCKPQELPPQCGEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLSKROEASRLDVCCCKPQELPPQCGEGLLGFFQKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNNSFTQKIQYATEFFEQSVLAIPASSWDDFDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNNSFTQKIQYATEFFEQSVLAIPASSWDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Qy 1021 NIKCPKGGGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGGGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140

Db 1081 RKVPGTDPAPFVPPYTTITNVFYEQYITILPEGLFMLSLCLVPTFAVSCILLGLDLRSGLL 1140
QY 1141 NLLSIWMLVDTVGFMALMDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIWMLVDTVGFMALMGISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVPAVAMTNLPGLVILGLAKAQLIQIPFFRLNLLITLLGLHLGLVFLPV 1260
Db 1201 AKEATISMSGSAVPAVAMTNLPGLVILGLAKAQLIQIPFFRLNLLITLLGLHLGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALQKRAEBAVAAMVASCNHPHSRVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVPALALQKRAEBAVAAMVASCNHPHSRVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 14

US-11-242-459-9
; Sequence 9, Application US/11242459
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-9

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGMLLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGMLLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLQKICERLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNP 120
Db 61 TPARKITGDHLILLQKICERLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNP 120
QY 121 VNLHCHTCSNQSLEFINTVRAQLGAGQLPAVVAEAFQHSFABQSQSDSCSRVRVPA 180
Db 121 VNLHCHTCSNQSLEFINTVRAQLGAGQLPAVVAEAFQHSFABQSQSDSCSRVRVPA 180
QY 181 ATLAVGTMCQVGSALCNARWLNFGQDGTGNGLAPLIDITFHLLPEQAVGSGIQIPLNEGV 240
Db 181 ATLAVGTMCQVGSALCNARWLNFGQDGTGNGLAPLIDITFHLLPEQAVGSGIQIPLNEGV 240
QY 241 ARCNEQGDVVATCSQDCQCAAPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVVATCSQDCQCAAPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSPNSOARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420

Db 361 IPVVALAAGLVFTELTTDPVELWSPNSOARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
QY 421 RYDSSLILGPKNPSGILDLLELELELQERLHLQVWSPQAORNISLQIDICYAPLNPNT 480
Db 421 RYDSSLILGPKNPSGILDLLELELELQERLHLQVWSPQAORNISLQIDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTPKDGKTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTPKDGKTALAL 540
QY 541 SCHADYGAPVFPFLAIGGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAOKLWEAELEEM 600
Db 541 SCHADYGAPVFPFLAIGGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAOKLWEAELEEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFPATSYIVIFLYISLALGSYSWSRV 660
Db 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFPATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATLGLGVAVVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGAADNIFIVLE 720
Db 661 MYDSKATLGLGVAVVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGAADNIFIVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEALCPFLGALTTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEALCPFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPPGQEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LEWITRGVVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFNRYFEVGPV 900
Db 841 LEWITRGVVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFNRYFEVGPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFTQKIQYATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFTQKIQYATEFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLCLKNCMSITMGSVRPSVEQPHKYLWPFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLCLKNCMSITMGSVRPSVEQPHKYLWPFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSQGLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVPPYTTITNVFYEQYITILPEGLFMLSLCLVPTFAVSCILLGLHLGLVFLPV 1140
Db 1081 RKVPGTDPAPFVPPYTTITNVFYEQYITILPEGLFMLSLCLVPTFAVSCILLGLHLGLVFLPV 1140
QY 1141 NLLSIWMLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIWMLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVPAVAMTNLPGLVILGLAKAQLIQIPFFRLNLLITLLGLHLGLVFLPV 1260
Db 1201 AKEATISMSGSAVPAVAMTNLPGLVILGLAKAQLIQIPFFRLNLLITLLGLHLGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALQKRAEBAVAAMVASCNHPHSRVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVPALALQKRAEBAVAAMVASCNHPHSRVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 15

PCT-US01-04098A-1830
; Sequence 1830, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides


```
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; TYPE: PRT
; LENGTH: 1359
; ORGANISM: Homo sapiens
PCT-US01-04098A-1830

Query Match          99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy      1  MAEAGLRGWLWALLRLRLAQSBPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
Db      1  MAEAGLRGWLWALLRLRLAQSBPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60

Qy      61  TPARKITGDHLLILLQKICRPLYTGPNTOACCSAKQLVLSLEASISITKALLTTCPCASDNF 120
Db      61  TPARKITGDHLLILLQKICRPLYTGPNTOACCSAKQLVLSLEASISITKALLTTCPCASDNF 120

Qy      121  VNLHCHNTCSPNQSLFINTRVAQLGAGOLPAVVAEAFYQHSFABQSDYSCSRVRVPA 180
Db      121  VNLHCHNTCSPNQSLFINTRVAQLGAGOLPAVVAEAFYQHSFABQSDYSCSRVRVPA 180

Qy      181  ATLAVGTMCGVYSALCNQRLWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db      181  ATLAVGTMCGVYSALCNQRLWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240

Qy      241  ARCNESQDDVATCSCQDCAACSPAIAEQALDSTFYLGMPGSLVLIILCSVFAVVTI 300
Db      241  ARCNESQDDVATCSCQDCAACSPAIAEQALDSTFYLGMPGSLVLIILCSVFAVVTI 300

Qy      301  LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWASWPLTILVLSV 360
Db      301  LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWASWPLTILVLSV 360

Qy      361  IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
Db      361  IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420

Qy      421  RYDSSLGPKNFGSILDLILLLELOERLRLQVMSPEAQRNISLQDICYAPLNPNDT 480
Db      421  RYDSSLGPKNFGSILDLILLLELOERLRLQVMSPEAQRNISLQDICYAPLNPNDT 480

Qy      481  SLVDCCINSLLQVFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL 540
Db      481  SLVDCCINSLLQVFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL 540

Qy      541  SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFTSLNNYPAGDPRLAQAKLWBEAFLEEM 600
Db      541  SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFTSLNNYPAGDPRLAQAKLWBEAFLEEM 600
```

```
Qy      601  RAFORMMAGMFQVFTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSRSRV 660
Db      601  RAFORMMAGMFQVFTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSRSRV 660

Qy      661  MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSLVLQVVPFLVLVSGADNIFIFYLE 720
Db      661  MVDKATLGLGGVAVVLGAVMAAMGFFSYLGRSLVLQVVPFLVLVSGADNIFIFYLE 720

Qy      721  YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSAICFFLGLATPMPAVRTFALTSLGLAV 780
Db      721  YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSAICFFLGLATPMPAVRTFALTSLGLAV 780

Qy      781  ILDFELLQMSAFVALLSLDSKQEAASRLDVCCKVQPELPPPGQEGEGLLGFQKAYAPFL 840
Db      781  ILDFELLQMSAFVALLSLDSKQEAASRLDVCCKVQPELPPPGQEGEGLLGFQKAYAPFL 840

Qy      841  LHWITRGVVLFLALFGVSLYSMCHISVGLDQBELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db      841  LHWITRGVVLFLALFGVSLYSMCHISVGLDQBELALPKDSYLLDYFLFLNRYFEVGAPV 900

Qy      901  YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATERPEQSYLAIPASSWVDDFDW 960
Db      901  YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATERPEQSYLAIPASSWVDDFDW 960

Qy      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP 1020

Qy      1021  NIKCPKGLAAAYSTSVNLTSDGQVLI-----ASRPMAYH 1053
Db      1021  NIKCPKGLAAAYSTSVNLTSDGQVLDVAILSPRLEYSGTISAHCNLYLLDLSASRPMAYH 1080

Qy      1054  KPLKNSQDYTEALRAARELAANITADLRKVPCTOPAREVPPVTIINVPEQVLTILPGL 1113
Db      1081  KPLKNSQDYTEALRAARELAANITADLRKVPCTOPAREVPPVTIINVPEQVLTILPGL 1140

Qy      1114  FMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGFMALWDISYNAVSLINLVS 1173
Db      1141  FMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGFMALWGISYNAVSLINLVS 1200

Qy      1174  AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db      1201  AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260

Qy      1234  LIQIEFFRLNLLITLLGLLHGVFLPVLISYVGPDPVNPALAEQKRAEAAVAAMVVASCP 1293
Db      1261  LIQIEFFRLNLLITLLGLLHGVFLPVLISYVGPDPVNPALAEQKRAEAAVAAMVVASCP 1320

Qy      1294  NHPSRVSTADNITYVNHSPSGSIKGAGAINFLPNNGRQF 1332
Db      1321  NHPSRVSTADNITYVNHSPSGSIKGAGAINFLPNNGRQF 1359
```

Search completed: April 7, 2006, 19:32:35
Job time : 576 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:23:20 ; Search time 77 Seconds
(without alignments)
2380.515 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAI NLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 584368 seqs, 137612332 residues

Total number of hits satisfying chosen parameters: 584368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New.*
1: /SIDSS/prodata/1/paa/pct_NEW_COMB.pep.*
2: /SIDSS/prodata/1/paa/US06_NEW_COMB.pep.*
3: /SIDSS/prodata/1/paa/US07_NEW_COMB.pep.*
4: /SIDSS/prodata/1/paa/US08_NEW_COMB.pep.*
5: /SIDSS/prodata/1/paa/US09_NEW_COMB.pep.*
6: /SIDSS/prodata/1/paa/US10_NEW_COMB.pep.*
7: /SIDSS/prodata/1/paa/US11_NEW_COMB.pep.*
8: /SIDSS/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	6	US-11-301-094-6
2	6872.5	99.5	1359	6	US-11-270-796-22
3	5407	78.3	1333	6	US-11-270-796-3
4	2402.5	34.8	1278	1	PCT-US06-05584-822
5	2402.5	34.8	1278	6	US-11-191-274A-129
6	2402.5	34.8	1278	6	US-11-191-274A-130
7	2402.5	34.8	1278	7	US-11-385-692-2400
8	2402.5	34.8	1278	7	US-11-385-692-2401
9	1341.5	19.4	1383	6	US-11-301-094-2
10	1046	15.1	1274	6	US-11-301-094-4
11	782.5	11.3	419	6	US-10-953-349-21065
12	588.5	8.5	967	6	US-10-461-673-16739
13	583.5	8.4	1182	6	US-11-332-764-2
14	583.5	8.4	1182	6	US-11-337-244-149
15	482	7.0	891	8	US-10-772-265-1197
16	455	6.6	1358	8	US-60-772-265-241
17	437	6.3	831	7	US-11-360-355-133805
18	349	5.1	783	6	US-10-461-673-16750
19	325	4.7	204	7	US-11-360-355-141351
20	261	3.8	575	7	US-11-360-355-120785
21	252	3.6	505	6	US-10-461-673-16871
22	229.5	3.3	465	7	US-11-360-355-141367
23	225	3.3	542	6	US-11-214-063A-1670
24	222.5	3.2	422	7	US-11-360-355-141365
25	219	3.2	1137	6	US-10-461-673-16889

26	219	3.2	1330	6	US-10-461-673-10541	Sequence 10541, A
27	190	2.8	126	7	US-11-360-355-141352	Sequence 141352, A
28	152.5	2.2	632	6	US-11-293-697-3548	Sequence 3548, Ap
29	141	2.0	164	7	US-11-360-355-127069	Sequence 127069, A
30	139.5	2.0	484	7	US-11-360-355-120169	Sequence 120169, A
31	135.5	2.0	252	7	US-11-360-355-140206	Sequence 140206, A
32	126	1.8	808	6	US-10-724-972B-7050	Sequence 7050, Ap
33	123.5	1.8	171	7	US-11-360-355-149589	Sequence 149589, A
34	121.5	1.8	1043	6	US-10-536-606-20	Sequence 20, Appl
35	120.5	1.7	2426	6	US-11-203-806A-11	Sequence 11, Appl
36	118.5	1.7	768	6	US-11-214-063A-2044	Sequence 2044, Ap
37	118.5	1.7	788	6	US-11-214-063A-1692	Sequence 1692, Ap
38	117.5	1.7	619	8	US-60-732-162-1828	Sequence 1828, Ap
39	116.5	1.7	619	6	US-11-312-958-46	Sequence 46, Appl
40	116.5	1.7	619	6	US-10-461-673-12167	Sequence 12167, A
41	114	1.7	488	8	US-60-752-355-45206	Sequence 45206, A
42	114	1.7	985	6	US-11-293-697-2874	Sequence 2874, Ap
43	112.5	1.6	697	6	US-10-703-799B-226	Sequence 226, App
44	111	1.6	201	7	US-11-360-355-152528	Sequence 152528, A
45	111	1.6	958	8	US-60-781-953-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-11-301-094-6
; Sequence 6, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-094-6

Query Match	100.0%	Score	6909;	DB	6;	Length	1332;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	1332;	Conservative	0;	Mismatches	0;	Indels	0;
		Gaps	0;				
QY	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60				
DB	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60				
QY	61	TPARKITGDHILILQKICPRLYTGNTQACSAQQLVSLASLSITKALLTRCPACSDNF	120				
DB	61	TPARKITGDHILILQKICPRLYTGNTQACSAQQLVSLASLSITKALLTRCPACSDNF	120				
QY	121	VNLCHNTCSNQSILFINVTRVAQIGAGQLPAAVYAEFYOHSPABQSDYSCSRVRPAA	180				
DB	121	VNLCHNTCSNQSILFINVTRVAQIGAGQLPAAVYAEFYOHSPABQSDYSCSRVRPAA	180				
QY	181	ATLAVGTWCGYVYGALCNQORWLNFGQDGTGNGLAPLDITFHLHPQAVGSGIQPLNEGV	240				
DB	181	ATLAVGTWCGYVYGALCNQORWLNFGQDGTGNGLAPLDITFHLHPQAVGSGIQPLNEGV	240				
QY	241	ARCNESQDDVATCSQDCAASCPAIPALPQALDSTFYLGQMPGSLVLIILICSVFVVTTI	300				
DB	241	ARCNESQDDVATCSQDCAASCPAIPALPQALDSTFYLGQMPGSLVLIILICSVFVVTTI	300				
QY	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPFQGHGTWASWPLTILVLSV	360				
DB	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPFQGHGTWASWPLTILVLSV	360				

```
QY 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQPPFRTNQVILLTAPNRSSY 420
DB 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQPPFRTNQVILLTAPNRSSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELELQERLRHLQVMSPEAQNISLQDICYAPLNPDNT 480
DB 421 RYDLSLLGPKNFGSLDLDLLELELQERLRHLQVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGTLAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGTLAL 540
QY 541 SCMDYGPAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPLAOKLWEAEFLBEM 600
DB 541 SCMDYGPAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPLAOKLWEAEFLBEM 600
QY 601 RAFQRMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
QY 661 MYDYSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDYSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQSGGLLGGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQSGGLLGGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVPGAPV 900
DB 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVPGAPV 900
QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNFSFTOKIOYATEFFPQSYLAIPASSWDDFDIW 960
DB 901 YFVTTLGYNFSSEAGNNAICSSAGCNFSFTOKIOYATEFFPQSYLAIPASSWDDFDIW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVTLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVTLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEPFPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEPFPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVPAGVAMTNLPGLVILGLAKAQLQIIFPFRNLTLTLGLHLGVFLPV 1260
DB 1201 AKEATISMGSAVPAGVAMTNLPGLVILGLAKAQLQIIFPFRNLTLTLGLHLGVFLPV 1260
QY 1261 ILSYVGPVNPALALQKRAEAAVAVMVASCPNHPSTRYSTADNIYVNHSPSGSIKGAGA 1320
DB 1261 ILSYVGPVNPALALQKRAEAAVAVMVASCPNHPSTRYSTADNIYVNHSPSGSIKGAGA 1320
QY 1321 ISNLFNNGRQF 1332
DB 1321 ISNLFNNGRQF 1332
```

RESULT 2

US-11-270-796-22

; Sequence 22, Application US/11270796

; GENERAL INFORMATION:

```
; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; CURRENT FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 22
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-270-796-22
```

```
Query Match 99.5%; Score 6872.5; DB 6; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
```

```
QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVNCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVNCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNQSLFINVTRVAOLGAGOLPAVVAYEAFYQHSFAEYSDSCSRVRVPA 180
DB 121 VNLCHNTCSNQSLFINVTRVAOLGAGOLPAVVAYEAFYQHSFAEYSDSCSRVRVPA 180
QY 181 ATLAGVTCMGVYGALCNAQRLNFGQDGTGNGLAPLDTITPHLLBPGQAVGSGIOPLNEGV 240
DB 181 ATLAGVTCMGVYGALCNAQRLNFGQDGTGNGLAPLDTITPHLLBPGQAVGSGIOPLNEGV 240
QY 241 ARCNEQDQDDVATCSCDCAAPALPQALDSTFYLGQMPGSLVLIILICSVFVVTTI 300
DB 241 ARCNEQDQDDVATCSCDCAAPALPQALDSTFYLGQMPGSLVLIILICSVFVVTTI 300
QY 301 LLVGFVRVAPADKSKMVDPKKGTSLSDKLSSTHTLLGQFFQGWGTWVSWPLTLTLVLSV 360
DB 301 LLVGFVRVAPADKSKMVDPKKGTSLSDKLSSTHTLLGQFFQGWGTWVSWPLTLTLVLSV 360
QY 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQPPFRTNQVILLTAPNRSSY 420
DB 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQPPFRTNQVILLTAPNRSSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELELQERLRHLQVMSPEAQNISLQDICYAPLNPDNT 480
DB 421 RYDLSLLGPKNFGSLDLDLLELELQERLRHLQVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGTLAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGTLAL 540
QY 541 SCMDYGPAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPLAOKLWEAEFLBEM 600
DB 541 SCMDYGPAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPLAOKLWEAEFLBEM 600
QY 601 RAFQRMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
QY 661 MYDYSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDYSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQSGGLLGGFFQKAYAPFL 840
```

Db 781 ILDFLLQMSAFVALLSLDSKQREASRLDCCVCKPQELPPQGGEGLLGFTQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGPV 900
Qy 901 YFVTTILGYNFSSEAGMNAICSSAGCNGNFSTQKIQYATEPFPQSYLAIPASSWVDDFDW 960
Db 901 YFVTTILGYNFSSEAGMNAICSSAGCNGNFSTQKIQYATEPFPQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDPR 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDPR 1020
Qy 1021 NIKCPKGGILAAVSTSVNLTSDDGVL-----ASRPMAYH 1053
Db 1021 NIKCPKGGILAAVSTSVNLTSDDGVLDTVAILSPREYSCTISAHNCNLYLLDSASRPMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYITINVFYEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYITINVFYEQYLTILPEGL 1140
Qy 1114 FMLSICLVPTFAVSCILGLDLSGLLNLISIVMILVDTVGFMAWDISYNAVSLINLYS 1173
Db 1141 FMLSICLVPTFAVSCILGLDLSGLLNLISIVMILVDTVGFMAWDISYNAVSLINLYS 1200
Qy 1174 AVGMSVEFVSHITRSPAIKTPWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAIKTPWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIPFFRLNLITLGLLHGLVFLPVILSYVGPVDPNPALEQKREABEAAVAVASCP 1293
Db 1261 LIQIPFFRLNLITLGLLHGLVFLPVILSYVGPVDPNPALEQKREABEAAVAVASCP 1320
Qy 1294 NHPSRVSTADNIYVNHSPGSGIKGAGAINFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPGSGIKGAGAINFLPNNGRQF 1359

RESULT 3
US-11-270-796-3
; Sequence 3, Application US/11270796
; GENERAL INFORMATION:
; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; TITLE OF INVENTION: RESTORATION OR MIMICRY PP p16 Ink4a ACTIVITY
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; PRIOR FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 3
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-270-796-3

Query Match 78.3%; Score 5407; DB 6; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

Qy 4 AGLRGWLLWALLRLQAQSPYTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSNTPA 63
Db 3 AAWQGWLLWALLNSAQGLYTPTHKAGCTFYECGKNPELSGSLMTLSNVCSLSNTPA 62
Qy 64 RKTGDHLLLOKICPRLYTGN-TOACCSAKOLYSLEASLSITKALLTRCPACSDNFVN 122
Db 63 RHVTGDHLLORVCPRLYNGENDTYACCSTKQLVSLDSSLSTIKALLTRCPACSDNFVS 122
Qy 123 LHCHNTCSPQSLFINVTRVAGLQAGQLPAVVAYEAFYQHSFABQSDYSCSRVRPAAAT 182

Db 123 IHCHNTCSPDQSLFINVTRVQDQQLPAVVAYEAFYQHSFABQSDYSCSRVRIPAAAS 182
Qy 183 LAVGTMCVGYGSAICNAORWLNFGQDTGNGLAPLDITPHLLBPGQAVGSGIOPNBSGVAR 242
Db 183 LAVGTMCVGYGSAICNAORWLNFGQDTGNGLAPLDITPHLLBPGQAVGSGIOPNBSGVAR 242
Qy 243 CNESQGDVATCSCDCAASCPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTILL 302
Db 243 CNESQGEDSAACSCDCAASCPIVPPPPALRPSFTYMGMRPGLALIIITPAVFLLSVVL 302
Qy 303 VGFRVAPARDKSNWDPKKGTSLSKLSFSTITLGOFPQCHGTWVASHPLILVLSVIP 362
Db 303 VYLRVASNRNKTKTAGSQEAPNLPRKRFPSPHTVLGRFPESWGTRVAVSPLVTLASFTV 362
Qy 363 VVALAAGLVFTLTDPVWELWAPNSQARSSEKAFHDOHGFPPFRFTNOVLTATPNRSSYV 422
Db 363 VTALS SVGLTFTLTDPVWELWAPNSQARSSEKAFHDEHGFPPFRFTNOIIVTAKNRSSYK 422
Qy 423 DSSLILGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 482
Db 423 DSSLILGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 482
Qy 483 YDCCINSILQYPPNNRTLLLTANOTLWGQTSQVDWKHFLYCANAPLTFKGTALALSC 542
Db 483 TDCCVNSLLQYPPNNHTLLLTANOTLWGQTSQVDWKHFLYCANAPLTFKGTALALSC 542
Qy 543 MADYCAPVFPFLAIGSYGKDYSEARALIMTFLSNYPAGDPRLAQAKLWBEAFLEEMRA 602
Db 543 IADYCAPVFPFLAIGSYGKDYSEARALIMTFLSNYPAGDPRLAQAKLWBEAFLEEMRA 602
Qy 603 FORMMAGMFQVFTTAERSLEDEINRTTAEDLPATSYIVIFLYISLALSGYSSSRVAV 662
Db 603 FORSTADKQIAFSAERSLEDEINRTTQDLFPVFAISYLVIFLYISLALSGYSSSRVAV 662
Qy 663 DSKATILGCGVAVVILGAVMAAGPFSYGLRSILVILQVVPFLVLSVGDNIPIFVLEYQ 722
Db 663 DSKATILGCGVAVVILGAVMAAGPFSYGLVSPSSVLIQVVPFLVLAAGADNIPIFVLEYQ 722
Qy 723 RLPRPMPGPREVHIGRALGRVAPSMILCSLSAICFFELGALTMPAVRTPALTSGLAIVL 782
Db 723 RLPRPMPGPREVHIGRALGRVAPSMILCSLSAICFFELGALTMPAVRTPALTSGLAIVL 782
Qy 783 DFLQMSAFVALLSLDSKQREASRLDCCVCKPQELPPQGGEGLLGFTQKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQREASRLDCCVCKPSSRLPPEKQEGLLGFTQKAYAPFLH 842
Qy 843 WITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGPV 902
Db 843 REIRPVWLLFLVLFGANLYLNCNISVGLDQDLALPKDSYLLDYFLFLNRYFVGPV 902
Qy 903 VTTILGYNFSSEAGMNAICSSAGCNGNFSTQKIQYATEPFPQSYLAIPASSWVDDFDW 962
Db 903 DTTSGYNFSTEAGMNAICSSAGCNGNFSTQKIQYATEPFPQSYLAIPASSWVDDFDW 962
Qy 963 P-SSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDPRN 1021
Db 963 PSSCCRLYITRGPHKDBFCPTDTSFNCLKNCNMTLGPVRPTTQPHKYLPWFNDPRN 1022
Qy 1022 IKCPKGGILAAVSTSVNLTSDDGVLASRFMAYHKLKNSQDYTEALRAARELAANITADLR 1081
Db 1023 IRCPKGGILAAVSTSVNLSDGQIIASQFMAYHKLKNSQDYTEALRAARELAANITADLR 1082
Qy 1082 KVPGTDPAPFVFPYITINVFYEQYLTILPEGLFMLSICLVPTFAVSCILGLDLSGLL 1141
Db 1083 KVPGTDPAPFVFPYITINVFYEQYLTILPEGLFMLSICLVPTFAVSCILGLDLSGLL 1142
Qy 1142 LLSIVMILVDTVGFMAWDISYNAVSLINLYSVAAGMSVEFVSHITRSPAIKTPWLER 1201
Db 1143 LLSIIMLVDTIIGLMAWGISYNAVSLINLYSVAAGMSVEFVSHITRSPAIKTPWLER 1202
Qy 1202 KEATISMGSAVAGVAMTNLPGILVLGLAKAQIPIFFRLNLITLGLLHGLVLPVI 1261
Db 1261 KEATISMGSAVAGVAMTNLPGILVLGLAKAQIPIFFRLNLITLGLLHGLVLPVI 1261

Db 1203 KDATIFMGSANFAGVAMTNFPGILLGPAQALQIIPFFRLNLLITLLGLLGLVFLPV 1262
Qy 1262 LSYGPDVNPALALOKRAEAAVAMVASCNHPSPRSTADNTYVNHSPGSG-IKGAGA 1320
Db 1263 LSYLPGDVNQALVLEKLATEA-AMVSBPSCPQYPPADANTSDYVNYGFNPEIPEINA 1321
Qy 1321 ISNFLPNNGROF 1332
Db 1322 ASSSLPKSDQKF 1333

RESULT 4
PCT-US06-05584-822
; Sequence 822, Application PC/TUS0605584
; GENERAL INFORMATION:
; APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION,
; TITLE OF INVENTION: ASSESSMENT, PREVENTION, AND THERAPY OF CANCER
; FILE REFERENCE: DPG-064.25(25992-6425)
; CURRENT APPLICATION NUMBER: PCT/US06/05584
; CURRENT FILING DATE: 2006-03-02
; PRIOR APPLICATION NUMBER: 60/690,064
; PRIOR FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: 60/654,227
; PRIOR FILING DATE: 2005-02-17
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 822
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US06-05584-822

Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLMALLRLRLAQSEPYTHIHPGFCAPYDECG-----KNPELSGSLMTLSNVCSLS 59
Db 4 RGLALGALLLLCPAQFSQ-----SCVMYGEIGYAKRNCYSEYSG-----46

Qy 60 NTPARKITGDHLLTLQKICPLRYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 47 --PPKELPKDGYDLVDELCEGFFFG-NVSLCCDVRLQTLKONLQLPLQFLSRCSFCYN 103

Qy 120 FVNHLCHNTCSNQSILFINVTR----VAQLGAGQLPAVVAEYAFYQHSFAESGYSDCSRV 175
Db 104 LMLFCELTCSPRQSQFLNVATATEDYVDPVNTQNTKTVKELQYVYVQGSFANAMYNACRDV 163

Qy 176 RVPAATLAVGTMCGYVGSALCNAQRWLPFGQDTGNGLAPLDIT-----PHLLBPQA 228
Db 164 EAPSSNDKALGLLCKGDADA-CNATWIEYMFKNKGQAPFTITPVDFPFW-----215

Qy 229 VSGSIOPLNEGVARCNESQDDVATCSQDCAACPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESDEVTAPECSQDCSIVCGPKQPPPPAPWILGLDAMYVM 272

Qy 280 QMPGSLVLIILCSVFAVWITLL-----VGRFVAPARKSKMVDPKKGTSL 326
Db 273 WITVMAFLVFFGAFFAVMCYRKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHLLGQFFQGWGTWVASWPLTILVSLVPIVVALAAGLVFTELTDPELWSAP 386
Db 325 DPVSAAFEGCLRLFTRWGSCFVRNPGCVIFFSLVFIACSSGLVFRVVTNPDVLSAP 384

Qy 387 NSOARSKAFDHFQGFRTNOVILTPARNSSRYRSDLLGPK-NFSGILDLLDLELL 445
Db 385 SSQARLEKEVDFHGFGEFFTEQLIRAPLTDKHIQYPSGADVPFGPLDIQLHQVL 444

Qy 446 ELQRLRLHLOWSPAEARQNSLQDICVAPLNPDNTSLYDCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLDQICLAPLSPYNT---NCTILSVLNYFQNSHVLDDHK 499

Qy 506 NQTLMGQTSQVMDKDHFLYCANAPLTKPGDTALALSCMADYDGPVFPFLPFIAGGYKGDYS 565
Db 500 GDDFF---VYADYTHFLYCVRAPASLNDTSLLDHPCLTGFGGVPFWLVLGGYDDQYN 556
Qy 566 EAEALIMTFSLNNYPAGDPRLAQAQKLEBAFLEEMRAFORMAGMFOVTFTAERSLDEI 625
Db 557 NATALVITFPVNNYNTDEKLRQAQWKEKFINVKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSYIVIFLYISALSGYSWSVMVDSKATLGLGGVAVVLGAVMAAMG 685
Db 614 NRESDDVFTTVISYALMFYISALGHIKSCRLLDVDSKVSIGIAGLILVSSVACSLG 673
Qy 686 FFSYLGRSILVILQVVPFLVLSVADNIFIFVLEYQRLPRRPPGEPREHVHTRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIPFLVAVGVNIFILVQAYQORDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLCSLSEACFFLIGALTMPAVRTTALISGLAVILDFLLOMSAFVALLSLSKQOAS 805
Db 734 SMFLSSPSETVAFELGALSVMPAVHTFSLFAGLAVFIDFLLOITCFVSLGLDIKQSKN 793
Qy 806 RLDVCCCKPOELPPGQ-GEGLLGPFOKAYAPFLHWTIRGVVLLFLALFVSLYLM 864
Db 794 RLDIFCCVRGAEDGTSSVQASECLFRFFKNSYSPULLKQWNPVIAIFVGLVFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYVVTTLGYNFSSEAGNAICSSAG 924
Db 854 NKVDIGLQSLMPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKQNMVCGMG 913
Qy 925 CNPFSFTKIQIYATEFPQSYLAIPASSWDDFDLWLT-SSCCELYISGNKDKFCPST 983
Db 914 CNNDSLVQOI FNAQAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSRVPSVEQFKYLPWFNDRNRIKCPKGLLAAYSTSVN--LTS 1040
Db 971 VDPACVR-CRPLTPEGKORPQGGDFRFLPMFLSDNPNKCGKGGHAAYSANVILGH 1029
Qy 1041 DQVQLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPQDTPAFEVPTITNV 1100
Db 1030 GTRVGATYFMTVHTVLTQTSADFIDALKKARLIASNV-ETWINGS--AYRVFPYSVFV 1086
Qy 1101 FVEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSTVMTILVDTVGPMLWD 1160
Db 1087 FVEQVLTITDITFNLGVSGLAIFLVTVMLGCELWSAVIMCATAMVLMVNMFGVWLMG 1146
Qy 1161 ISYNAVSLINLSVAGMSVEFVSHITRSFAISTKTWLERAKEATISMSAYFAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIPFFRLNLLITLLGLLHGLVFLPVILSYVGPDVNPA 1272
Db 1207 FGIIVVLAPAKSQIPIQIFFRMYLAMVLLGATHGLIFLPLVLLSYIGPSYNKA 1258

RESULT 5
US-11-191-274A-129
; Sequence 129, Application US/11191274A
; GENERAL INFORMATION:
; APPLICANT: Applera Corporation
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001536PROV
; CURRENT APPLICATION NUMBER: US/11/191,274A
; CURRENT FILING DATE: 2005-07-28
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-191-274A-129

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;

Matches	522;	Conservative	237;	Mismatches	450;	Indels	103;	Gaps	25;
Qy	7	RGWLLMALLRLAQSBBPYTHIHPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS	59						
Db	4	RGLALGLLLLLLCPAQVFSQ-----SCWYIGECGIAYGDKRYNCEYSG-----	46						
Qy	60	NPARKITGDHLLILLQKICPLRYTGPNTYQACSSAKQLVSLSEASISITKALLTRCPACSDN	119						
Db	47	--PPKLPKPGDYDLVQLCLCPGFFG-NVSLCCDVRQLQTLKDNILQLFLQLSRCPSPFYN	103						
Qy	120	FVNLLHCHNTCSPNOSLFINVTR-----VAQLGAGQLPAVVAEAFVQHSFAEQSDSCSRV	175						
Db	104	LLNLFCBLTCSPROQFLNVTATEDYDVPVNTQKTNVKELYYYVGQSFANMAYNACRDV	163						
Qy	176	RVPAATAVLAVGTMGVYGSALCNAQRWLNFQDGTGNGLAPLDIT-----PHLLEPGA	228						
Db	164	EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKNDGQAPPTITPVDFSDPVH-----	215						
Qy	229	VGSGIQPLNEGVARCNESQGDVVATCSQDCQCAASCPAIAEPQ-----ALDSTPYLG	279						
Db	216	---GMEPMNNAATKGCDSVDEVTAPCSQCDCSIVCGPKPQPPPPPPAPWTILGLDAMTVIM	272						
Qy	280	QMPGSLVLIILCSVEAVVITLL-----VGRVAPARDKSKMVDPKCKGTSLS	326						
Db	273	WITYMAFLLVFFGAFVAVWCYRKRYFVSEYTPIDSNIAFSV-NASDRGE-----ASCC	324						
Qy	327	DKLSFSYTHLLGQPFQCGWGTWASWPIITLVSVIPVVALAAGLVFPELATTDPVELWSAP	386						
Db	325	DPVSAAPGEGCLRLFTWGSFCVNPCCVIFFSLSVFTACSSGLVFRVVTNPVDLWSAP	384						
Qy	387	NSQARSKAFHDQHGFPPRTNQVILTPANRSSYRYDSLLGPK-NPFGILDLDLLELL	445						
Db	385	SSQARLEKBYFDQHFQGFPRTEQLIIRAPLTDKHIYQYPYPSGADVPPGPLDIQLHQVL	444						
Qy	446	ELQERLHLQWSPEAQRNISLODICVAPLNPNQTSLYDCINSLLOYFQNNRTILLTLTA	505						
Db	445	DLQIAIEN---ITASYDNETVLDICIAUPSYPNT---NCTILSVLMYFNQNSHVDLHKK	499						
Qy	506	NOTLMGQTSQVDWKDHFLYCANAPLTPKDGTLALSCWADYGAPVFPFLAIGGYKGDYS	565						
Db	500	GDDEF---VYADVHTHFLYCVRAPASLNDTSLLDHPCLTGTFGGVPFWLVGGYDQVYN	556						
Qy	566	EAEALIMTFSLNNYPAGDPRLAQAKLWEBAFLBEMRAFORRMAGMFQVITFAERSLEDEI	625						
Db	557	NATALVITPVNNYNDTEKLQRAQAWKEKFINPVKNYKN---PNLTISPTAERSIEDEL	613						
Qy	626	NRTAEDLPATNSYIVIFLYISLAGSYSSWRGVMWDSKATLGLGVAVVLGAVMAAMG	685						
Db	614	NRESDSVFTWVISYATIMFYIISLAGLHIKSCRRLLVDKSVLSGIAGILVLSVACSLG	673						
Qy	686	FFSYLGRSSILVILQVVPFLVLSVGADNIIFVLVEYQRLPRRGPPEPREVHIGRAIGHVAP	745						
Db	674	VFSYIGLPLTILVIEVLPFLVLAVGVNIIIFILVQAYORDERLOGETUQQGLRVLGSEVAP	733						
Qy	746	SMILCSLSEAIACFFGLALTMPAVRTPALTSGLAVILDFLLQMSAFVALLSLDSKQREAS	805						
Db	734	SMFLSSPSETVAFPLGALSVMPAVHTSFLAGLAVFIDFLQITCFVSLGLGLDKRQKN	793						
Qy	806	RLDVCCCKQKQELPPPPQ-GEGLLGFQKAYAPFLHLWITRGVWLILLFLALFGVSLYSYM	864						
Db	794	RLDIFCCVRGAEDGTSVQASESCLREFRFKNSYSPLLLKDMWRPIVIAIFVGVLSFSIAVL	853						
Qy	865	CHLSVGLDQELALPKDSYLLDYFLFNRYFEVGAQPVYFVTTLGVNPFSSSEAGMAVICSAG	924						
Db	854	NKVDIGLDQSLSPDDSDIMVDYFKSISQYLHAGPPVYFLVEEGHDYTSKQNMVCGMG	913						
Qy	925	CNNFSPQTKIQATPEPPEQSYLAIPASPSVDDDFDMLTLP--SSCCRLYISGPNKDKCFPST	983						
Db	914	CNNDLSLQQIFNAQDLNRYTRIGFAPSSWIDDYFDWKVQPOSSCCRV---DNITDQFCNAS	970						
Qy	984	VNSLNLCKNCMSIT-MGSVRPSVEQFHKYLPLWFILNDRPNITCKPKGLAAYSTSVN--LTS	1040						
Db	971	VDPAPCYR-CRPLTPEGKQFQGGDFMRFLPMLFSLNDNPNKPCKGKGGHAAVSSAVITLLGH	1029						

RESULT 6					
US-11-191-274A-130					
; Sequence 130, Application US/11191274A					
; GENERAL INFORMATION:					
; APPLICANT: Applera Corporation					
; APPLICANT: Bruno DOMO					
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF					
; FILE REFERENCE: CU001536PROV					
; CURRENT APPLICATION NUMBER: US/11/191,274A					
; CURRENT FILING DATE: 2005-07-28					
; NUMBER OF SEQ ID NOS: 334					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 130					
; LENGTH: 1278					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-11-191-274A-130					
Query Match 34.8%; Score 2402.5; DB 6; Length 1278;					
Best Local Similarity 39.8%; Pred. No. 1.5e-185;					
Matches 52; Conservative 237; Mismatches 450; Indels 103; Gaps 25;					
Qy	7	RGMLLWALLRLAQSBPYTTHIHPGGYCAFVDECG-----KNPELSGSLMTLSNVSCLS	59		
Dd	4	RGLAGLLLLLLCPAQVFSG-----SCVMYGEGCIAYGDKRYNCEYVG-----	46		
Qy	60	NTPARKITGDHILLXKI CPLYTGNTQAACSAKQLVSLEASLTSTKALTTRCPACSDN	119		
Dd	47	--PPKLPKDGYDLVQELCPGFFFG--NVSLCCDVRLQTLLKONLQLPLQFLSRCPSCFVN	103		
Qy	120	FVNLRHCNTCSPNQSIPI NVTR----VAQLGAGQLPAVVAYEAFYOHSPAQSYSDCSRV	175		
Dd	104	LMLPCLRLTCSPRQS QFLNVATATEDYVDPTNQTNNVKELQYYVQGSPANAMYNACRDV	163		
Qy	176	RVAAAAATLVGTGCMGVYGYSALCNARWLNFQDGDTGNGLAPLDIT-----PHILLEPGQA	228		
Dd	164	EAFSSNDKALGLLCGRKADA-CNAWNIEWFNKGNGQAPPFITPVFSDFPVH-----	215		
Qy	229	VGSIGIPLMEGVARNCHESOGDDVATCSQCDAASCIPAIRPO-----ALDSTFYLG	279		
Dd	216	---GMEPMNNA TKGCDESDEVTA PCSCQDCSIVCGPKQP PPPPPAPWTILGLDAMYVM	272		
Qy	280	QMPSGLVLIIILCSFAAVTILL-----VGRVPAPARDKS KMVDPKKGTSL S	326		
Dd	273	WIYMAFLIVFGAFP VAWCYRKRYFVBSEYTPIDSNIAFSV-NASDKGE-----ASCC	324		
Qy	327	DKLSFTHTTL LGQFP CGWGTVASWPLTI LVISVPVVALAA GLVFTELTTDPVELWSAP	386		
Dd	325	DPVSAAFEGLRLRF TRWGSFCVRNP GCVI FF SLVP ITACSGL VFVRTVTNPVDLWSAP	384		
Qy	387	NSQAAREKA FHDOHGCFPRTR TQVILTAPNR SYRDSL LLGP K -NFSG ILDLLLELL	445		
Dd	385	SSQAARKEGYFD HGGFFFRTEQL IIRAP LT DKHI YQYPVS GADVF PGFPDI QLHVLA	444		
Qy	446	ELQERLRHLQVMS PAQRNISLODIC YAPLNPDNTSL YDCCINS LILOYPN NRTLILLTA	505		

Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVNLVFNQNSVLDHKK 499
Qy 506 NQTLMGQTSQVMDKDFHLYCANAPLTFKDGFTALALSCNADYCAPVFPFLAIGGKYGKDY 565
Db 500 GDDPF--VYADYHTFLYCYRAPASLNDTSLHDPCLTGTPGVFPLVGLGGYDDQNYN 556
Qy 566 EBAELIMFTSLNNYPAGDPLRAQAKLWEEAFLEEMRAFQRMAGMFWQVTFPFAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEKFNFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPIFATSYIVIFLYISLALSGSYSSWSRVMVDSKATLGLGGVAVLGVAMAMG 685
Db 614 NRESDSDFTVVVISYAIMFLYISLALGHIKSCRRLLVDSKVSGLIAGLILVLSVACSLG 673
Qy 686 PFSYIGIRSSILVLOQVFPFLVSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VPSYIGLPLTLVIEVIFPLVLAGVDNIFILVQAYQDERLQGETLQQLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICTFFLGALTPMPAVTFFALTSGVLVLOFLOMSAFVALLSDSKRQEAS 805
Db 734 SMPLSFSFSETVAFFLGALSVMPAVTFFSLFAGLAVFIDFLQITCFVSLGLDKEQEK 793
Qy 806 RLDVCCCVKQBELPPPGQ--GEGLLGFFQKAYAPFLHWHITRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASECLFRFFKNSYSPLLLDKDMRPVIAIFVGLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTLTLYNFSSEAGNNAICSSAG 924
Db 854 NKVDIGLQDLSMPDDSYMVDYFKSISQYLHAGPPVYFVLEBGHDYTSKQGMVCGMG 913
Qy 925 CNNFSFTQKIQVATFPPEQSYLAIPASSWDDFIDWLTP--SSCCRLYISGPNKDFCPST 983
Db 914 CNNDLSVQOIFNAALQDNTYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR--CRPLTPGKQPOGGDFMRFLPFLSNDPNPKCGKGHAAYSSAVMILLGH 1029
Qy 1041 DGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPFGTDPAPFVFPYITNV 1100
Db 1030 GTFVGATYFTVHTVTLQTSADFDALKKARLIASNT--ETMGINGS--AYRVFPYSVFV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLPPTFAVSCLLGLDLBSGLNLLSIVMILVDTVGFMAWD 1160
Db 1087 FVEQYLTIIDDTIFNLGVSIGAIFLVTMVLGCELMSAVIMCATIAMVLVNNFVGMWLMG 1146
Qy 1161 ISUNAVSLNLSAVGMSVEFVSHITRSPAISTKPTMLERAKEATISMGSAVPAGVAMTN 1220
Db 1147 ISUNAVSLNLSVSCGISVEFCSHITRATVSMKGRVEREALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVLGAKAQIIOIFFRMLLTLGLHLGLVFLPVILSVYGVDPNPA 1272
Db 1207 FGGIVVLAFKSIQIFQIFFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVYKA 1258

RESULT 7
US-11-385-692-2400
; Sequence 2400, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385, 692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 2400
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2400

Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
Qy 7 RGMILLMALLRLAQSEPYTHIHQPGYCAFVDECG-----KNPELSGSLMTLNVSCLS 59
Db 4 RGLAELGLLLLLCPAQVFSQ-----SCVMYGECCGAIYDKRYNCEYSG-----46
Qy 60 NTPARKITGTHLILLOKICPRLYTGPNTQACCSAKOLVSLASISITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVBLQCPGFFFG--NVSCLCDVRQQLKONLQLPQFUSRCPCSFYN 103
Qy 120 FVNHLCHNTCSNOSLFINVTR-----VAQLGAGQLPAVAYEAFYQHSFABQSDSCSRV 175
Db 104 LILNLFCELTCSPROSQFLNVTATEDYDVPVNTQKTNVKELOYYVVGGSFANAMYNACRDV 163
Qy 176 RVPAAATLAVGTMCGVYGSALCNAQRLNFGDGTGNGLAPLDIT-----PHLLEPQQA 228
Db 164 EAPSSNDKALGLLCKGDADA--CNATNMEYMFNKGNGOAPFTITPVSDFPVH-----215
Qy 229 VGSIGIPLNEGVARNESQGDVATCSODCAASCAPAIARPQ-----ALDSTFVLG 279
Db 216 ---GMEPMNNAATKGDSEVDVETAPCSQDCSIVCGPKPQPPPPAPWTIILGLDAMYIM 272
Qy 280 QMPGSLVLIILCSVFAVVTILL-----VGFVRVAPARDKSMVDPKKGTSL 326
Db 273 WITWMAFLVFFGFAVWCYRKYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLQGFQGGTWSAWPLTILVSVIPVVALAAGLVFTLTTDTPVELWAP 386
Db 325 DPVSAAPFGLRLRFLTRWGSFCVRNPGCVIEFSLVFTACSSGLVFRVVTNPVDLWAP 384
Qy 387 NSQARSEKAFHDQHGPFERNQVLTAPNRSSVRYDSLLLGPK--NFGSILDLDLLLELL 445
Db 385 SSQARLEKEYFDQHGPFRTQELIIRAPLTDKXHYQYPGSGADVFPGPPLDIQILHQVL 444
Qy 446 ELQERLRHLQVMSPEAQRNISLQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVNLVFNQNSVLDHKK 499
Qy 506 NQTLMGQTSQVMDKDFHLYCANAPLTFKDGFTALALSCNADYCAPVFPFLAIGGKYGKDY 565
Db 500 GDDPF--VYADYHTFLYCYRAPASLNDTSLHDPCLTGTPGVFPLVGLGGYDDQNYN 556
Qy 566 EBAELIMFTSLNNYPAGDPLRAQAKLWEEAFLEEMRAFQRMAGMFWQVTFPFAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEKFNFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPIFATSYIVIFLYISLALSGSYSSWSRVMVDSKATLGLGGVAVLGVAMAMG 685
Db 614 NRESDSDFTVVVISYAIMFLYISLALGHIKSCRRLLVDSKVSGLIAGLILVLSVACSLG 673
Qy 686 PFSYIGIRSSILVLOQVFPFLVSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VPSYIGLPLTLVIEVIFPLVLAGVDNIFILVQAYQDERLQGETLQQLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICTFFLGALTPMPAVTFFALTSGVLVLOFLOMSAFVALLSDSKRQEAS 805
Db 734 SMPLSFSFSETVAFFLGALSVMPAVTFFSLFAGLAVFIDFLQITCFVSLGLDKEQEK 793
Qy 806 RLDVCCCVKQBELPPPGQ--GEGLLGFFQKAYAPFLHWHITRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASECLFRFFKNSYSPLLLDKDMRPVIAIFVGLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTLTLYNFSSEAGNNAICSSAG 924
Db 854 NKVDIGLQDLSMPDDSYMVDYFKSISQYLHAGPPVYFVLEBGHDYTSKQGMVCGMG 913
Qy 925 CNNFSFTQKIQVATFPPEQSYLAIPASSWDDFIDWLTP--SSCCRLYISGPNKDFCPST 983
Db 914 CNNDLSVQOIFNAALQDNTYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040

```
Db 971 VDPACVR-CRPLTPGKORPOGGDFMFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMYTHVLTQTSADFALKARLIASNT-ETMGINGS--AYRVPYSVFVY 1086
Qy 1101 FYEQYLTILPEGLFMLSLCLVPTFAVSCILLGLDLRSGLNLLSIYMLIVDTVGMALWD 1160
Db 1087 FYEQYLTIIIDDTIFNLGSLGALFLVTMVLGCELWSAVIMCATIAMVLVNMFGVMWLG 1146
Qy 1161 ISYNAVSLNLVSAVGMSEVFSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKGSVEREABEALAHMGSSVPSGILTJK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRMLLITLLGLHLGFLVPLVILSYVGPVNP 1272
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLANVLLGATHGLIFLPLVILSYIGPSVYKA 1258

RESULT 8
US-11-385-692-2401
; Sequence 2401, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385,692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2401
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2401

Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLAQSEPYTHIQPGYCAPYDRCG-----KNPELSGSLMTLSNVCLUS 59
Db 4 RGLALGLLLLLCPAQVFSQ-----SCWYGEAGIAYGDKRYNCEYSG----- 46

Qy 60 NTPARKITGDHLLILQKICPRELYTGNTQACCSAKQLVSLASLSITTKALLTRCPACSDN 119
Db 47 ---PPKPLPKDGYDLVQELCPGFFG--NVLSCDVRQLQTLKONLQLPLQFLSRCPSCFYN 103

Qy 120 FVNLHCHNTCSPNQSLFINVTR---VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCELTCSPRSQSOFLNVATEDYDVPVTWQTKNVKELQYVYQSPFANNAYNACRDV 163

Qy 176 RYPAATLAVGTWCGYVYGALCNAQRWLNFGQDGTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCCGKQADA-CNATWIEYMFENKONGQAPFTITPVFSDFPVH----- 215

Qy 229 VCSGTOPLNEGVARNESQGDVATCSODCAASCPAIAAPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESDEVTAPCSQDCSIVCGPKQPQPPPPAPMTWILGLDAMYIM 272

Qy 280 QMPGSLVLIILCSFVAVVTILL-----VGRFVAPARDKSKWDPKKGTSLS 326
Db 273 WITYMAFLVFFGAPFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHTLLGQFFQGWGTWASVWMLTTLVLVSVIPWVALAGLVFTLTDPVBLWSAP 386
Db 325 DPVSAFEGCLRLFRWGSFCVRNPGCVIFPSLVFTACSSGLVFRVTRTNPVFLWSAP 384

Qy 387 NSQARSEKAFPHQHPPTTQVILTAPNRSYYSYDLSLLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYPQHFGPPFTTQILRAPLTDKHLYQPYPSGADVPFPGPLDIQLHQL 444
```

```
Qy 446 ELQERLRHLQVWSPEAQRNLSIQDICYAPINPDNTSLYDCINSLSLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTIQLDCLAPLSPNT---NCTILSVLNYFQNSHVDLHKX 499
Qy 506 NOTLMGQTSQVDKWHFLYCANAPLTFKQGTALALUSCMADYDAPVPPFLAIGSKYKQDYS 565
Db 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLHDPLCLGTFCGVPVFWLVLGGYDQNTN 556
Qy 566 EBAALIMTFSLNNYPAGDPRLAQAKLWBEAPLEEMRAFORRMAGMPQVFTTFAERSLEDEI 625
Db 557 NATALVIFPVNNYNDTEKQRAQAWKEKFINFVKNYKN---PNLTISFTAERSLEDEL 613
Qy 626 NRTTAEDELIPATSYIVIFYISLALGSYSSRNVMSKATILGGVAVLIGVAVMAAMG 685
Db 614 NRESDSDFVTWVISVAINFLYISLALGHKSKRLLVDSKVSIGIAGILVLSSVACSLG 673
Qy 686 FESYLIGIRSSLVILQVPELVLSVGDNITFIFVLFYQRLPRPRGPREVHIGRALGRVP 745
Db 674 VFSYIGLPLTLIVIEVPELVLAGVDNIFILVQAYQDRERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLGEAICFFLGCALTPEPAVRTFALTGSLAVILDLPLOWSAFVALLSDSKRQEAS 805
Db 734 SMFLSFSETVAFFLGALSVMPEAVTFFSLFAGLAVFIDFLQLTFCVSLGLDIDKQKKN 793
Qy 806 RLDVCCCKVPQBLPPPGQ--GEGLLLGFFQKAYAPFLLHWITRGVLLLLFLALFGVLSYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFPKNSVSPILLKXMMRPVIAIFGVLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLNRYFVGAPVYFVTTLYGYNFSSBAGNAICSSAG 924
Db 854 NKVDIGLDQSLSMDDSYMDVYFKSISQYLHAGPPVYFVLEBHDYDTSKQGNMVCGGMG 913
Qy 925 CNNEPFTOKIQVATEFPBQSYLAIPASSWVDDFIDWLTG--SSCCRLYISGPNKDCPST 983
Db 914 CNDSLVQLQIFMAAQLDNTRIGPAPSNWIDIDYDWWKVPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLCKNCSIT--WGSVRPSVEQFHKYLPWFLNDRPNIKCPKGGLAAYTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPGKORPOGGDFMFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMYTHVLTQTSADFALKARLIASNT-ETMGINGS--AYRVPYSVFVY 1086
Qy 1101 FYEQYLTILPEGLFMLSLCLVPTFAVSCILLGLDLRSGLNLLSIYMLIVDTVGMALWD 1160
Db 1087 FYEQYLTIIIDDTIFNLGSLGALFLVTMVLGCELWSAVIMCATIAMVLVNMFGVMWLG 1146
Qy 1161 ISYNAVSLNLVSAVGMSEVFSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKGSVEREABEALAHMGSSVPSGILTJK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRMLLITLLGLHLGFLVPLVILSYVGPVNP 1272
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLANVLLGATHGLIFLPLVILSYIGPSVYKA 1258

RESULT 9
US-11-301-094-2
; Sequence 2, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 1383
```



```
Db 284 LWFILKYSCHRSPNGEGCTVDLKGK-----LEVFEGLCARYANAVTK 330
Qy 351 WPLTILVLSPVVALAAG-LVFETLTDPELWAPNSQARSEKAF-HDQHFGEFFRTN 408
Db 331 HPLIFVSLGLIVAAACCSGNFKFSLTHSDVQVSAADGETRNEKKFIHS--FGPNHLE 388
Qy 409 QUILTAPNRSSRYDLSLLGPKNFSGILDLDLLELLELOERLHUVQWSPRAQRNLSIQ 468
Db 389 QIFINLPPTT-----KSMFNPLFEEMFQLVGNQILNT--ACYGNSSVKLD 432
Qy 469 DICYAPLNNTSLYDCCINSLLQYFQNNRTLLLTANQTLTMCQTSQVDWKDHPLYCANA 528
Db 433 DICYPIGKKN-----HGCALMSPNTYFQNKWTNFENAGPPTIDDEIFDDQHWELKYCIRN 488
Qy 529 PLTFKDGDTALALSCMADYGAPVFPFLAIGG-----YKGDQ-YSEAEALIMTFSLNPNYPAGD 583
Db 489 PLTV--STYSEMSCGEGSPDIPILVFGSNESIKGAEMYTYTARTIMITVLIRG----- 541
Qy 584 PRLAOKLWEAPLEMRAPFORMMAGFOVTTAERSLEDEINRTTAED--LPIFATSYI 641
Db 542 PE-DOAIAMETAPLNMMSYKMHANF---TPMTETSAVEIHTAVETDKIVSVIACAIV 597
Qy 642 VFLYISLALGSY--SSSRVWVDSKATLGLGVAVVLGAVMAAGFFSYLGIIRSLVIL 699
Db 598 LIWITMLGINHWPESSILSALVHHKLLISISAVMISVISVWCISGMFSLFGVHATDNAI 657
Qy 700 QVVPFLVLSVGADNIPFVLEYQRLPRRPGEP-----REVH--IGRALGVAPSMLLCSLS 753
Db 658 VVLPFVITCLGINRPIVIRITQANGCHYGLPNISYRENNHRI SNVRRSIPVLTNSLI 717
Qy 754 EACIFPL-GALTP-----MPAVRTFALTSLGLAVILDFLQMSAFVALSLDSKQ----- 802
Db 718 CSTCLFLAGGLVLPYVSVPVAVFARHAGLAILMDTAFYLLVMLPLFQYDARREMSGRK 777
Qy 803 -----EASRLDVCCVKQELPPPCQGGELLGPFQKAYAPFLHWTIRGVVLLP 853
Db 778 EIWPMYELNESKINCLMEAVDGNLRSP-----VDWFKLAIAPLLKKICRIWIATPF 830
Qy 854 LALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLPLNRYFVGAPVYVPTTLGYNFSSE 913
Db 831 FVSLIACYCTCLGFGNQWAFSETSYLTKHQFQNMENINIGPLWFFVSGDVKNHDP 890
Qy 914 AGMAICSSAGCNNSFTQKIQ---YATFPFQSYLAIPASSWVDDFDLWLTTP--SSCRL 969
Db 891 KMOKFCTLAGCDNDSMGNKIRSLAENY-KGNYLHGDVNIWLDLYLQFMHPRGSCCKM 949
Qy 970 YISGPNKDFC-RSTVNSLNCNKMSITMGSVRPSVEQFHKYLPWFLNDRNPNIKCPKGG 1028
Db 950 -----DGKQFCDDPS--NATHC-SSCSSVASLTTTEYEFYRNLHHFLETPPSIQCAHGG 1001
Qy 1029 LAAYSTSVNLTSDGOVLASRFMAYHKP--LKNSQDYTEALRAARELANITADLRKVPT 1086
Db 1002 MALAKPAINLTKNGIKQSYFTFPKMLNSIQIYDARFAKYLADDIEREL-EIPGV 1060
Qy 1087 DPAFEPYPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLLSTV 1146
Db 1061 ----KVYVYSTFPFYEQYLTSTTYTLVLVLVFAVFTISLFLRVNLAGSLVTVFVLL 1116
Qy 1147 MLVDVTGFMALWDISYNAVSLINLVSAGMVEFVSHITRSFAISTKPTWLERAKBATI 1206
Db 1117 SSYLHLMEMMYLLGITVNVSVINMAMSLGIAVEFFGQMLHGFFYNSKKPKREERAPALV 1176
Qy 1207 SMGSVAFVAGVAMTNLPGILV----LGLAKAQIQLIFFEFLNLALTLLGLHLGLVPLV 1262
Db 1177 SNGSTLSGI-----FPAINITAGLSFADSRVLITTFPCNQLVGLVCAVHGVVYMTLL 1232
Qy 1263 SYVGPDVNPALAEQKRAEEA 1283
Db 1233 AIFGSDFYQNVSEESTDEA 1253
```

RESULT 11

```
US-10-953-349-21065
; Sequence 21065, Application US/10953349
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 2004-09-30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21065
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-953-349-21065
```

```
Query Match 11.3%; Score 782.5; DB 6; Length 419;
Best Local Similarity 38.4%; Pred. No. 7.1e-55;
Matches 164; Conservative 88; Mismatches 138; Indels 37; Gaps 10;

Qy 863 SMC-HISVGLDQELALPKDSYLLDYFLNRYFVGAPVYVPTTLGYNFSSEA-GMAIIC 920
Db 2 ALCTRIEAGLEQIALLPRDSYLGQYFSNISEYLRVGPPLYFVVK-DYNSLESKHTNQLC 60
Qy 921 SSAGCNNSFTQKIQYATEFPFQSYLAIPASSWVDDFDLWLTTPS--SCRLYISGPNKDK 978
Db 61 SISHCDNSLLNEISRASLVPTSSVIAPKPAASWLDLFWISPEAFSCCKP----TNDS 116
Qy 979 FCPSTVNSLNCNKMSITMGSV-----RPSVEQFHKYLPWFLNDRPNI 1022
Db 117 YCPDDQPPCCLDPEGCGLGVCCKDTTCFRHSDLVNDRPSTAQPREKLPWFLDALPSA 176
Qy 1023 CKPKGGLAAYSTSVNLT--SDGOVLASRFMAYHKPNSQDYTEALRAARELANITADL 1080
Db 177 DCAKGHGAATNSVDLNGVEGVIQASEFTYHTPLNRQDVTYNAIRAARDFAIISSL 236
Qy 1081 RKVPGTDPAFEPYPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 237 K-----MDIPFYSVFIPEQYLDIWKALINITVALGAIFV-CLITSSVWSSAI 287
Qy 1141 NLLSTVMLVDVTGFMALWDISYNAVSLINLVSAGMVEFVSHITRSFAISTKPTWLER 1200
Db 288 LLLVLMIILDLGMVMAILQILNNAVSVNVNLTMSIGIAVEFCVHIVHAFMWSIGDR-SQR 346
Qy 1201 AKEATISGSAVPAGVAMTNLPGILVGLAKAQIQLIFFEFLNLALTLLGLHLGLVPLV 1260
Db 347 AKTALCTMGASVFGITLTKLVGLVLCFSTSEIFVYVYFQMYLALVIIIFLHLGLVPLV 406
Qy 1261 ILSVYVGP 1267
Db 407 VLSLFGP 413
```

```
RESULT 12
US-10-461-673-16739
; Sequence 16739, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Kyle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
```

```
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yunging
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 16739
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-461-673-16739

Query Match      8.5%; Score 588.5; DB 6; Length 967;
Best Local Similarity 21.3%; Pred. No. 1.4e-38;
Matches 216; Conservative 192; Mismatches 359; Indels 249; Gaps 37;

QY 334 HT-----LLGQFQGGTWWASWPLTILVSVIPVVALAAGLVF--TELTTPDVELWSAP 386
DB 115 HTDCLGGLSRTFQWLQWQGAHPWIFLLAPLMTAALGTGFLYLPKDEEEDLEEHTYTV 174
QY 387 NSQARSEKAFHDHFGFFFTNQVILTAPNRSS-----YRYDSLILGPKNPSGIL 436
DB 175 GSPAKAERRPVQGH-----FTNDSYRFSASRRSTEANFVSLLVVSYSDSLDDPATFAVS 230
QY 437 DLDLLELLELBQLRHLQWSPQARNISLQDIC--YAPLPNDNTSLYDCCT--NSLLQ 492
DB 231 KLDGAVQDLRV-----AREKGSQIQYQVCARYAL-----CVPPNPILV 270
QY 493 YFQNNRTLLLTANTQTLMGQTSQVDWKDHFLYCANAPLTFKGTALALSCMADYGA-PVP 551
DB 271 AWQVNTKTLN-----SSISFPAY-----NHGRHPLY 296
QY 552 PFLAIGY-----KGKDYSEABALIMTFLSNYPAGDPRL-AQAKLMEAEFLBEMRAF 603
DB 297 LTGFFGGYILGSLGMLGQLLRKAKMLLYLK---TEDEYDVQSKQWMLTHLDDQFTNI 353
QY 604 QRMA-----GMFQVTFPAERSLEDEINRTTAEDLPFATSYVIFVYISL 649
DB 354 KNILAKKIEVPGVGVLGGQEKVWHFTSLRQLEPEATSVTVIPVPHLAYILILFAVT 413
QY 650 ALGSYSWSRWVDSKATLGLGGVAVVLGAVMAAMGFFSVLGRSSLVILQVVPFLVSV 709
DB 414 SCRFD-----CIRNKNCAVAFGISAFVLAWSGFGLLHIGV-PFVIIIVANSFPFILGV 467
QY 710 GADNIFIFVLEYQRLPRPGEPRVHIGRALGRVAPSMLLCSLSEAICFFLGALTMPAPV 769
```

```
DB 468 GVDDFMISAWHK-TNLGDIIE-RMSNVYSKAAVSITITITITNIALALYTGIMSSFRSV 525
QY 770 RTFALTSGVLDFLLQMSAFVALLSLDSKQESRLDVCCKVPQELP-----819
DB 526 QCFCIYTGTTLLFCYFYNIITCFGAFMALDGREV-----VCLCWLKKAADPKWPFKFC 580
QY 820 -----PGQCEGL--LLGPFQKAYAPFLHWHTRGVVLLFLALFGVSLYSMCHISVGL 871
DB 581 PFGSVDEHGTDDHPMSLFRDYFGFLTRSESKYFVFIYVLYIISIIYGCFFHVGSL 640
QY 872 DOELALPKDSYLLDYFLFLNRYF-EVGAPVYFVTLGYNFSSEAGMNAICSSAGCNFSF 930
DB 641 DLRLASDSDSYITPYFNVEENYFSDYGRVMVITVKVDYWDK-----DV 685
QY 931 TQKIQYATEFPEQSYLAIP--ASSWVDDFDIWLFPSSCCRLYISGPNKDKFCPSTVNSLN 988
DB 686 RQKLENTCTKIPKKNVYVDKNTLTFWLDAYVQ-----YLGKNSOD---PNEKNT-- 730
QY 989 CLKNCMSITMGSVRPSVEQPHKYLPLWFLNDRPNIKCPKGLAAYSTSVNLTSDGQVLASR 1048
DB 731 -----GLFMLSICLVP-----TFVSCLLGLDLRLSGLLNLISVILV 1150
QY 1109 LPE-----GLFMLSICLVP-----TFVSCLLGLDLRLSGLLNLISVILV 1150
DB 810 LEDTVRNVLVASAAMFIVSLLLIPYPLCMLVTFAGSVIVGV-----852
QY 1151 DTVGPMALWDISYNVSLINLVSAVMSVGFVSHITRSFAISTKPTWLERAKEATISMGS 1210
DB 853 --TGFWAFKYNLDSISMINLVICGFSDFAHSIYAFVSSQSPVQKSVQKSVQKSVQKSV 910
QY 1211 AVFAGVAMTNLPGLVLGLAKAQLIQIFFERLNLITLLGLLHGLVFLVPLVLSYVG 1266
DB 911 PVLQS-AISTTIGVCVLAARAKAYIPRT-FFKIMFLVMIFGAHGLIFIPVFLTFPG 964

RESULT 13
US-11-332-764-2
; Sequence 2, Application US/11332764
; GENERAL INFORMATION:
; APPLICANT: Wisotzkey, Robert G.
; TITLE OF INVENTION: PATCHED HOMOLOG 2 (PTCH-2) DISRUPTIONS, COMPOSITIONS AND METHODS
; FILE REFERENCE: R1673 CIP/75658.064000
; CURRENT APPLICATION NUMBER: US/11/332,764
; CURRENT FILING DATE: 2006-01-13
; PRIOR APPLICATION NUMBER: US 60/413,543
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 10/669,143
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mus musculus
; US-11-332-764-2

Query Match      8.4%; Score 583.5; DB 6; Length 1182;
Best Local Similarity 24.2%; Pred. No. 4.7e-38;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SFSTHTLLGQ-----FFQGS-----WGTWASWPLTILVSVIPVVALAAGLVTFELT 376
DB 20 SSAPHILAGSLQAPLMLRAYFQGLLPSLGRICKGKGVFLGLVAFGALALGRVAVIE 79
QY 377 TDPVELWSPNQSARSEKAFHQHFG-PFFRTNQVILTAPNRSSYRYSYDLSLLGPKNFSGI 435
DB 80 TDLEQLWVVGSRVSQELHYTKKUGEEAAYTSQMLI-----QTAHQEGGNVLTPE---A 131
```


Db 302 TARDLQQLRAEALQSTFLMSRQLYEHFRGDYQTHDIGNSEBQASMLVQAQMRFPVQ 361
QY 610 MPQVTTFAERSLE-----DEINRTAEDLPFATSYIV--IFLYISIALGYSYWS 658
Db 362 LAQEALPANASQOIHAFFSTTDLILRAFSE-----VSTRVGGYLLMLAYACVTMLRWD 417
QY 659 RVMDSKATLGGVAVVAGVMAAMGFSSVYLGRSSVLIVQVPPFLVSLVSGADNIFV 718
Db 418 --CAOSQAGVAGVLLVALAVASGLICALLGITFNATQVLPFLGALGVDDIFLLA 475
QY 719 LEYQRLPRRPGPREVHHIGRALGRVAPSMLLCSLSEACFFLGAALTMPAVRTFALTSL 778
Db 476 HAFTKAP--PDTPLRMEGCLRSTGTSTVNTVMVAFMAALVPALRAFSLQAAI 533
QY 779 AVILDFLOMSAFVALLSLSKQASRLDVCCC-----VKQOE----- 817
Db 534 VVGCNFAAVMLVFPAILSLDLRRRQRDLVLCFSPSCSAQVIOQLPQELGDRAPVGI 593
QY 818 -----LPP-----PG-----QEGG-- 826
Db 594 AHLTATVQAFTHCEASSQHVVTILPQAHLLSPASDPGLGSELYSGGSTRLDLSQEGTG 653
QY 827 -----LLLGFPQKAYAPFLHMITRGVWLLFLFALFGVSLYSMCHISVGLDQ 873
Db 654 POAACRPLLCAHWTLAHPARYQFAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713
QY 874 ELALPKOSVLLDYFLNRYFEVGAPVYVTTLGNFS--SEAGMAICSSACGNNFSTQ 932
Db 714 TDVPRGTKEHAFLSAQLRYFSL--YEVALVTQGGFDYAHSORAL-----FDLHQ 761
QY 933 KIQYATEFPEQSVLAIPAS-----SWDDF-----IDMLTPSSCCRLYISGPNK 976
Db 762 RFSSL-----KAVLPPTAQAPRTWLHYRSLQIOAFOQDQWASGRITCHSYNGSED 816
QY 977 D-----KFCPSTVNSLNCNCHSITM-----GSRVPSVEQPHKYLPMFLNDRPNKCPK 1026
Db 817 GALAYKLLIQTGNAQBPDPFSQLTTRKLVDEKGLIPP--ELFVMTGLTVWVSDPL----- 869
QY 1027 GGLAA-----YST-----SVNLTSDGOVLASRFWAYHKPLKNSQDYTEALR 1067
Db 870 --GLAASQANFYPPPEWLHDKYDVTTCENLRIPAAQPLEFAQFPFLHGLQKTAQDFEATE 928
QY 1068 ARELAANI--TADLRKVPGTDAPEFEPVYITNVFYEQLTILPEGLFMLSIC--LVPTF 1124
Db 929 GARAACTEAGQAGVHAYPSGSP-----LFWEQYGL--RECFLVACILLVCTF 976
QY 1125 AVSCLLLGLDLRGLNLLSIVMLVDTVGFMAWMDISYNVSLINLVSAGVMSVEFVSH 1184
Db 977 LVCALLLSPTAGLI--VLVAMMTVELFGIMGFLGILKLSAIPVVLVASIGIGVEFTVH 1035
QY 1185 ITRSFALSTKPTWLERAKEATISMGSAVFAGV---AWTNLPGLVLGLAKAQLIQIFFR 1241
Db 1036 VALGFUTSHGSRNLRAA-----SALRQTPAPVTDGAVSTLLGLLMLAGSNFDFIRYFFV 1090
QY 1242 LNLILLGLLHGLVFLPVILSVXP 1267
Db 1091 VLTVTLGLLHGLLIPVLLSILGP 1116

RESULT 15

US-60-772-265-1197

; Sequence 1197, Application US/60772265

; GENERAL INFORMATION:

; APPLICANT: BOUKHAROV, ANDREY A.

; APPLICANT: DU, ZIJIN

; APPLICANT: GUO, LIJANG

; APPLICANT: HRESKO, MICHELLE C

; APPLICANT: KOVALIC, DAVID K

; APPLICANT: ZHAOLONG, LI

; APPLICANT: LU, MAOLONG

; APPLICANT: MCCARTER, JAMES P

; APPLICANT: MILLER, NANCY M

; APPLICANT: VAUDIN, NANCY

; APPLICANT: WILLIAMS, DERYCK J
; APPLICANT: WU, WEI
; TITLE OF INVENTION: IDENTIFICATION AND USE OF TARGET GENES
; TITLE OF INVENTION: FOR CONTROL OF PLANT PARASITIC NEMATODES
; FILE REFERENCE: MNDI:002USP1
; CURRENT APPLICATION NUMBER: US/60/772,265
; FILING DATE: 2006-02-10
; NUMBER OF SEQ ID NOS: 1919
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1197
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: genomic DNA sequence=SeqID_396; coding sequence=SeqID_981
US-60-772-265-1197

Query Match 7.0%; Score 482; DB 8; Length 891;

Best Local Similarity 22.2%; Pred. No. 5,38,30;

Matches 203; Conservative 160; Mismatches 360; Indels 190; Gaps 35;

QY 540 LSCMADYCAPVPFPLAIGGYK---GKDYSEAEALIMTFSLNNYPAGD--PRLAQAK--L 591
Db 10 MNCPEP-----LIVGIRRDAGQVQIRHAEAFQTVFLVAS--AGDVQRLRSKNHR 59
QY 592 WEAEFLBEMRAFQRRMAGMFQVTFABRSLEDEINRTTAEDL----PIFATS----- 639
Db 60 MEKGVSTQAHAGATAAAMQORNT--KSIYDHTLNKPAQGIRVVHPPLASTSIQDMLRQF 117
QY 640 ----YVIFPL--YISLALGYSYSNRSV-----WVDSKATLGLGVAVVGLVMAAMGFF 687
Db 118 SEQFPVIFIGVLMII--YAGWSQVHWQGMFYSKSCLLAIIGLVITLASVAGLGUS 175
QY 688 SYLGIRSLVILQVVPRLVLSVGADNIFIFVLEYQRLPRRPGPREVHHIGRALGRVAPSM 747
Db 176 TAMNHFNAATQIVPFLTLGLGDDMFLHNYNDVLE--AVRQEVAVLLKETGMSV 232
QY 748 LLCLSSEACFFLGAALTMPAVRTFALTSLGLAVILDFLQLMSAFVALLSLSKQOEASRL 807
Db 233 LITSINNILAFUTGILPILALRSFCGQVAILLTSNVLCIILLFPALFALLDLRRKAGHR 292
QY 808 DVCCC-----VKQELPPQGE-----GALLGFFQKAVAP 838
Db 293 DMSFCSSRNQLVKDSTVTTHSSIEDLQMSAAVPTSSATKSLHKWTTVLGFLHGYIP 352
QY 839 FLHWHITGVVLLFLALFGVSLYSMCHISVGLDQELALPKDYSLLDYFLNRYFEVGA 898
Db 353 LLRRFIKTVLLVCAANFLFCGFLYYSRIGLELADVLPEHTAPAAFLKARERYFSP-Y 411
QY 899 PVYFY---TTLGYNFSSBAGMAICSSAGCNF-----SFTQKI 934
Db 412 PMFIVPKGPMVDY--ANQHQKLEQLRARDIGRSNFVVKVDGEPQWMMFMRTWLHLSQSSL 470
QY 935 QVATE--FPQSYLAIPASSWV-----DPIDMLTPSSCCRLYISGNKDKFCPSTVN 985
Db 471 DRAAKIGIPDDDLKXMAADALARNFTLSDDFL--LARKLLC-----SAQO 514
QY 986 SLNC--LKNCMSITMGSVRPSVEQPHKYL--PMFLNDRPNKCPKG-----LAAYS 1033
Db 515 RENCTEMRGAKLIEDGRINP--RGFNVLTAFYQDNMMYVYSQAFFPTPLPWHSAA 572
QY 1034 TSNVLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEFV 1093
Db 573 ESWVPPAD--PLLYSQIPPFYMMGLTDTQSIVQMIKEIRAI CDYSAD-----GLPVY 622
QY 1094 PTTITNVFEQYLTILPEGLFMLSICLVPTFAVSCLLGLDLRLSGLNLSIWMI--LYD 1151
Db 623 PSGIPPTFEQYKLT---FYLVSAILIIGVAVLIVISIIIFNPWAAAMVAVIVTWVTE 679
QY 1152 TVGFMAWMDISYNVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSA 1211
Db 680 LAGFMGVGVGNPISAVTLITAVGIVEFTAVHVLAFITS-----LGRDRMRWACLEH 734

THIS PAGE IS BLANK

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:36:06 ; Search time 201 Seconds
(without alignments)
2911.704 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAI SNFLPNNGRQF 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_21.*

1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	8	ADJ27297 Human NPC
2	6909	100.0	1332	9	ADY60740 Human Nle
3	6909	100.0	1332	9	AEB93570 Human NPC
4	6896	99.8	1332	4	AAM79169 Human pro
5	6896	99.8	1332	4	AAG65638 SSD-conta
6	6872.5	99.5	1359	8	AAM79168 Human pro
7	6872.5	99.5	1359	8	ADJ27337 Human NPC
8	6872.5	99.5	1359	9	ADY60780 Human Nle
9	6872.5	99.5	1359	9	AEB93610 Human NPC
10	6536	94.6	1344	4	AEB22593 Novel hum
11	5421.5	78.5	1331	8	ADJ27295 Rat NPCIL
12	5421.5	78.5	1331	9	ADY60738 Rat Niema
13	5421.5	78.5	1331	9	AEB93568 Rat NPCIL
14	5407	78.3	1333	8	ADJ27305 Mouse NPC
15	5407	78.3	1333	9	ADY60748 Mouse Nle
16	5407	78.3	1333	9	AEB93578 Mouse NPC
17	4466	64.6	982	4	ABG22691 Novel hum
18	2402.5	34.8	1278	3	AAB88445 Human NPC
19	2402.5	34.8	1278	3	AAB42983 Human ORF
20	2402.5	34.8	1278	8	ADQ39879 Human myo
21	2402.5	34.8	1278	8	ADJ06723 Novel bro
22	2385	34.5	1319	2	AAW88446 Mouse NPC
23	2262	32.7	1287	4	AB61737 Drosophil
24	2262	32.7	1287	8	ADS96670 Drosophil

25	1864.5	27.0	1223	4	ABB58629	Abb58629 Drosophil
26	1329	19.2	1170	2	AAW88447	Aaw88447 Yeast NPC
27	1078	15.6	1296	2	AAW88448	Aaw88448 Caenorhab
28	963	13.9	229	4	AAM80153	Aam80153 Human pro
29	963	13.9	229	4	AAM80152	Aam80152 Human pro
30	942	13.6	194	4	AG65637	Ag65637 SSD domai
31	651.5	9.4	1447	2	AAR75375	Aar75375 Human pat
32	651.5	9.4	1447	2	AAW52200	Aaw52200 Human pat
33	651.5	9.4	1447	2	AAW72969	Aaw72969 Human pat
34	651.5	9.4	1447	4	AAB67163	Aab67163 Human pat
35	651.5	9.4	1447	5	AAE19830	Aae19830 Human pat
36	651.5	9.4	1447	5	ABU10931	Abu10931 TRC8 rela
37	651.5	9.4	1447	5	AAW79571	Aaw79571 Human pat
38	651.5	9.4	1447	7	ABU62275	Abu62275 Human pat
39	651.5	9.4	1447	7	ADD46678	Add46678 Human Pro
40	651.5	9.4	1447	7	AD894224	Ad894224 Human ptc
41	651.5	9.4	1447	7	ADH62731	Adh62731 Human pat
42	651.5	9.4	1447	8	AD848989	Ad848989 Human pat
43	646.5	9.4	1434	2	AAW52199	Aaw52199 Mouse pat
44	646.5	9.4	1434	2	AAW72968	Aaw72968 Mouse pat
45	646.5	9.4	1434	4	AAB67159	Aab67159 Murine pa
46	646.5	9.4	1434	5	AAW79572	Aaw79572 Mouse pat
47	646.5	9.4	1434	7	ABU62271	Abu62271 Mouse pat
48	646.5	9.4	1434	7	ADH62722	Adh62722 Mouse pat
49	646.5	9.4	1434	7	ADH62722	Adh62722 Mouse pat
50	646.5	9.4	1434	8	AD848980	Ad848980 Mouse pat
51	643.5	9.3	1434	2	AAR94380	Aar94380 Mouse pat
52	628	9.1	1296	7	AAW47157	Aaw47157 Nevoid ba
53	588.5	8.5	967	7	ADJ21144	Adj21144 Novel hum
54	588	8.5	954	5	ABB98519	Abb98519 Human tes
55	584	8.5	1203	5	AAW43261	Aaw43261 Human pat
56	584	8.5	1203	5	AAE19829	Aae19829 Human pat
57	584	8.5	1203	6	ABG74104	Abg74104 Human pat
58	583.5	8.4	1182	5	AAE19831	Aae19831 Mouse pat
59	583	8.4	1203	3	AAW92703	Aaw92703 Human pat
60	580	8.4	1203	2	AAW28444	Aaw28444 Human ptc
61	579	8.4	1146	3	AAW92225	Aaw92225 Human pat
62	531	7.7	1061	4	ABB65553	Abb65553 Drosophil
63	526.5	7.6	1311	2	AAW72971	Aaw72971 Precis co
64	526.5	7.6	1311	4	AAB67156	Aab67156 Butterfly
65	526.5	7.6	1311	5	AAW79573	Aaw79573 Butterfly
66	526.5	7.6	1311	7	ABU62149	Abu62149 Buckeye p
67	526.5	7.6	1311	7	AEB94209	Aeb94209 Peacock b
68	526.5	7.6	1311	7	ADH62716	Adh62716 Butterfly
69	526.5	7.6	1311	8	AD848974	Ad848974 Butterfly
70	514	7.4	1405	8	ADN22811	Adn22811 Bacterial
71	512.5	7.4	1311	2	AAW52197	Aaw52197 Precis co
72	474	6.9	1286	4	ABB59092	Abb59092 Drosophil
73	474	6.9	1286	5	ABJ10929	Abj10929 TRC8 rela
74	464.5	6.7	1299	2	AAW86304	Aaw86304 Drosophil
75	464.5	6.7	1299	4	AAE05389	Aae05389 Human pat
76	464.5	6.7	1299	4	AAW85751	Aaw85751 Hedgehog
77	464.5	6.7	1299	4	AAW85751	Aaw85751 Hedgehog
78	464.5	6.7	1299	5	AAW85751	Aaw85751 Hedgehog
79	464.5	6.7	1299	6	AD25309	Ad25309 Hedgehog-
80	464.5	6.7	1299	6	AD25309	Ad25309 Hedgehog-
81	464.5	6.7	1299	7	ABW00878	Abw00878 Drosophil
82	464.5	6.7	1299	7	ABW01685	Abw01685 Drosophil
83	464.5	6.7	1299	7	ADH61129	Adh61129 Drosophil
84	464.5	6.7	1299	7	ADH54692	Adh54692 Drosophil
85	464.5	6.7	1299	8	ADU26454	Adu26454 Drosophil
86	464.5	6.7	1299	9	ADW25903	Adw25903 Drosophil
87	464.5	6.7	1299	9	ADW25903	Adw25903 Drosophil
88	438.5	6.3	1285	2	AAW72372	Aaw72372 Drosophil
89	438.5	6.3	1285	4	AAW72372	Aaw72372 Drosophil
90	438.5	6.3	1285	5	AAW72372	Aaw72372 Drosophil
91	438.5	6.3	1285	7	ABU62150	Abu62150 Fruit fly
92	438.5	6.3	1285	7	AD894211	Ad894211 Fruit fly
93	438.5	6.3	1285	7	ADH62718	Adh62718 Fly patch
94	438.5	6.3	1285	8	ADH62718	Adh62718 Fly patch
95	379	5.5	714	8	ADN22757	Adn22757 Bacterial
96	369.5	5.3	803	8	ADQ66737	Adq66737 Novel hum
97	351.5	5.1	767	5	ABB98520	Abb98520 Human tes

98 349 5.1 783 7 ADJ21155
99 339 4.9 783 6 ADA54204
100 316.5 4.6 608 7 ADD46676

Ad21155 Novel hum
Ada54204 Human pro
Add46676 Rat Prote

ALIGNMENTS

RESULT 1

ADJ27297
ID ADJ27297 standard; protein; 1332 AA.

XX AC ADJ27297;
XX 20-MAY-2004 (first entry)
XX DE Human NPC1L1.

XX Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX OS Homo sapiens.
XX PN WO2004009772-A2.

XX 29-JAN-2004.

XX 17-JUL-2003; 2003WO-US022467.

XX 19-JUL-2002; 2002US-0397442P.

XX (SCHE) SCHERING CORP.

XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX WPI; 2004-132945/13.

XX N-PSDB; ADJ27296.

XX GENBANK; AF192522.

XX New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting
PT or identifying antagonists of NPC1L1 for inhibiting intestinal
PT cholesterol absorption in a subject, or for treating elevated serum
PT cholesterol or stroke.

XX Claim 11; SEQ ID NO 4; 125pp; English.

XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1
(NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains
CC a motif ADJ2731 which acts as a trans-golgi network to plasma membrane
CC transport signal, and which exhibits limited tissue distribution and
CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a
CC sterol regulated element binding protein 1 (SREBP1) binding consensus
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
CC sensing domain (SSD) which is involved in sensing cholesterol levels,
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1
CC has 42* amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.

XX SQ Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 8; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEBYTTTHOPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEBYTTTHOPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLILLQKICPLRYTGPNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPLRYTGPNTQACCSAKQLVSLASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNOSLFINTVTRVAQLGAGQLPAVVAEAFYQHSFAESYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNOSLFINTVTRVAQLGAGQLPAVVAEAFYQHSFAESYDSCSRVRPAA 180
Qy 181 ATLAAGTMCVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
Db 181 ATLAAGTMCVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSCODCAASCPAIARPOALDSTFYLGOMPGSLVLIILICSVFVVTTI 300
Db 241 ARCNEQGDVATCSCODCAASCPAIARPOALDSTFYLGOMPGSLVLIILICSVFVVTTI 300
Qy 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLGGQFGQGTWVASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLGGQFGQGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPVELMSAPNSQARSEKAFHDQHFPPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTPVELMSAPNSQARSEKAFHDQHFPPFRTNQVILTAPNRSSY 420
Qy 421 RYDSLLGLPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSLLGLPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNLSLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMSQTSQVDMKOHFLYCANAPLTFKDTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMSQTSQVDMKOHFLYCANAPLTFKDTALAL 540
Qy 541 SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Qy 601 RAFQRMAGMFQVTFPERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSWSRV 660
Db 601 RAFQRMAGMFQVTFPERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSWSRV 660
Qy 661 MVDSKATILGLGSAVVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720
Db 661 MVDSKATILGLGSAVVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPMPAVRTPALTSLAV 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPMPAVRTPALTSLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKPQELPPPCQGGELLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKPQELPPPCQGGELLGFFQKAYAPFL 840
Qy 841 LHMVTRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHMVTRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNGNFSFTQKIQVATFEPQSYLAIPASSWDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNGNFSFTQKIQVATFEPQSYLAIPASSWDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCCLKMCSITMGSVRPSVEQFHKYLIPWFNLDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCCLKMCSITMGSVRPSVEQFHKYLIPWFNLDRP 1020
Qy 1021 NTKCPKGLAAYSTSVNLTSDQVLA SRFMAYHKPLKNSQDYTEALRAARELAANTADL 1080
Db 1021 NTKCPKGLAAYSTSVNLTSDQVLA SRFMAYHKPLKNSQDYTEALRAARELAANTADL 1080

QY 1081 RKVPGTDPAPFPYPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db |||||
1081 RKVPGTDPAPFPYPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
Db |||||
1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVAGVAMTNLPGILVGLAKAQLIQIPFRLNLITLGLLHGLVFLPV 1260
Db |||||
1201 AKEATISMSGSAVAGVAMTNLPGILVGLAKAQLIQIPFRLNLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db |||||
1261 ILSYVGPDPVPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNELPNNGRQF 1332
Db |||||
1321 ISNELPNNGRQF 1332

RESULT 2
ADY60740
ID ADY60740 standard; protein; 1332 AA.
XX AC ADY60740;
XX XX
19-MAY-2005 (first entry)
DE Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 4.
XX XX
KW Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX OS Homo sapiens.
XX XX
PN WO2005015988-A1.
XX XX
PD 24-FEB-2005.
XX XX
PF 16-DEC-2003; 2003WO-US040113.
XX XX
PR 17-JUL-2003; 2003US-00621758.
PR 22-AUG-2003; 2003US-00646301.
PR 16-SEP-2003; 2003US-00663208.
XX XX
PA (SCHE) SCHERING CORP.
XX XX
PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2005-284403/29.
DR N-PSDB; ADY60739, ADY60787.
XX XX
PT Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
XX PS Claim 30; SEQ ID NO 4; 146pp; English.
XX XX
CC The invention relates to an isolated Niemann-Pick C1-like protein 1
CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,
CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
CC recombinant vector comprising the polynucleotide, a host cell comprising
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
CC produce any functional NPC1L1 protein), an offspring or progeny of the

CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
CC dosage form, and information indicating that NPC1L1 is a target of
CC ezetimibe or substituted azetidinone), decreasing the level of intestinal
CC sterol or Salpha-sterol absorption in a subject (involving reducing the
CC level of expression of NPC1L1 in the subject), identifying an antagonist
CC of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a
CC functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist
CC of NPC1L1 which is useful for inhibiting or decreasing the level of
CC NPC1L1 mediated sterol or Salpha-sterol uptake in a subject such as
CC mouse, rat or human. The level of expression of NPC1L1 in the subject is
CC reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is
CC useful for screening a sample for intestinal sterol or Salpha-sterol
CC absorption antagonist. The NPC1L1 antagonists identified are useful for
CC treating medical conditions caused or mediated by NPC1L1, e.g.,
CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
CC arteriosclerosis. The present sequence represents a mammalian NPC1L1
CC protein.
XX SQ Sequence 1332 AA;
Query Match 100.0%; Score 6909; DB 9; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLWGWLWALLRLAQLAQSEPTTTHQPGVCAFYDECGKNPGLSGSLMTLSNVSCLSN 60
Db 1 MAEAGLWGWLWALLRLAQLAQSEPTTTHQPGVCAFYDECGKNPGLSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLLOKI CPRLYTGPNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLLOKI CPRLYTGPNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNPQSLF INVTRVAQLGAGQLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSNPQSLF INVTRVAQLGAGQLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTWCGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLEPGQAVSGIQLNEGV 240
Db 181 ATLAVGTWCGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLEPGQAVSGIQLNEGV 240
QY 241 ARCNEQGDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVLEWLSAPNSQARSEKAFPHQHPFRFTNOVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTDPVLEWLSAPNSQARSEKAFPHQHPFRFTNOVILTAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLDDLLLELQERLHLQVWSPQAORNTSLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDDLLLELQERLHLQVWSPQAORNTSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGFTAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGFTAL 540
QY 541 SCMADYGAPVFPFLAIGYKGYKDYSEALINTFSLNNYPAGDPLAQAKLWEAFLEEM 600
Db 541 SCMADYGAPVFPFLAIGYKGYKDYSEALINTFSLNNYPAGDPLAQAKLWEAFLEEM 600
QY 601 RAFQRMAGMFOVTFATRSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSVR 660
Db 601 RAFQRMAGMFOVTFATRSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSVR 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSSLVLQVVPFLVLSGADNIFVLE 720
Db |||||

Db	661	MVDSKATLGLGGVAVVLGAVNMAAGFFSYLGIRSSLVILQVVPFLVLSVGDNIFFIVLE	720
Qy	721	YQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTTPMPAVRTFALTSLGLAV	780
Db	721	YQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTTPMPAVRTFALTSLGLAV	780
Qy	781	ILDFLLQWSAFVALLSLDSKQBSRLDYVCCCVKPOELPPQGEGLLLGFFQKAYAPFL	840
Db	781	ILDFLLQWSAFVALLSLDSKQBSRLDYVCCCVKPOELPPQGEGLLLGFFQKAYAPFL	840
Qy	841	LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGPV	900
Db	841	LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGPV	900
Qy	901	YFVTTLYGNFSSEAGNNAICSSAGCNFSFTQKIQYATEFFPEQSYLAIPASSWDDDFIDW	960
Db	901	YFVTTLYGNFSSEAGNNAICSSAGCNFSFTQKIQYATEFFPEQSYLAIPASSWDDDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDFCPSTVNSLNCMKMSITWGSVRPSVEQFHKYLPWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDFCPSTVNSLNCMKMSITWGSVRPSVEQFHKYLPWFLNDRP	1020
Qy	1021	NLKCPKGGIAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL	1080
Db	1021	NLKCPKGGIAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL	1080
Qy	1081	RKVPGTDPAPFVFPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL	1140
Db	1081	RKVPGTDPAPFVFPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL	1140
Qy	1141	NLLSIWMLVDTVGFMAWDISYNAVSLNLVSAVGMSEVFVSHITRSPAIKPTWLER	1200
Db	1141	NLLSIWMLVDTVGFMAWDISYNAVSLNLVSAVGMSEVFVSHITRSPAIKPTWLER	1200
Qy	1201	AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFRLNLITLGLLHGLVFLPV	1260
Db	1201	AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFRLNLITLGLLHGLVFLPV	1260
Qy	1261	ILSYGPDVNPALALQKRAEBAVAAMVAVSCPNSHPSRVSTADNIYVNHSPGSIKGAGA	1320
Db	1261	ILSYGPDVNPALALQKRAEBAVAAMVAVSCPNSHPSRVSTADNIYVNHSPGSIKGAGA	1320
Qy	1321	ISNFPNNGRQF	1332
Db	1321	ISNFPNNGRQF	1332
RESULT 3			
AEB93570			
ID	AEB93570	standard; protein; 1332 AA.	
AC	AEB93570;		
XX			
DT	06-OCT-2005	(first entry)	
XX			
DE	Human NPC1L1 protein.		
XX			
KW	Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;		
KW	cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;		
KW	arteriosclerosis; human.		
OS	Homo sapiens.		
XX			
PN	WO2005069900-A2.		
XX			
PD	04-AUG-2005.		
XX			
PF	14-JAN-2005; 2005WO-US001469.		
XX			
PR	16-JAN-2004; 2004US-0537341P.		
XX			
PA	(MERI) MERCK & CO INC.		
XX			

PI	Garcia-Calvo M;		
XX			
DR	WPI; 2005-564070/57.		
DR	N-PSDB; AEB93569.		
XX			
PT	Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by		
PT	contacting NPC1L1 with detectably labeled substituted 2-azetidinone		
PT	glucuronide and a candidate compound and determining if compound binds to		
PT	human NPC1L1.		
XX			
PS	Example 2; SEQ ID NO 4; 215pp; English.		
XX			
CC	The invention relates to identifying a ligand of NPC1L1. The method		
CC	involves contacting human NPC1L1 with a detectably labeled substituted 2-		
CC	azetidinone glucuronide and a candidate compound and determining whether		
CC	the candidate compound binds to human NPC1L1. In identifying a ligand of		
CC	NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10		
CC	nM or lower. The detectably labeled substituted 2-azetidinone glucuronide		
CC	is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands		
CC	are useful for stimulating or blocking the activity of NPC1L1, and for		
CC	treating conditions caused or mediated by NPC1L1. It is useful for		
CC	reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart		
CC	disease, stroke, or arteriosclerosis. The present sequence represents a		
CC	human NPC1L1, a N-glycosylated protein.		
XX			
SQ	Sequence 1332 AA;		
Query Match 100.0%; Score 6909; DB 9; Length 1332;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAEAGLRGWLWALLRLLAQSEPYTTTHQPGYCAFYDECGRNPGLSGSLMTLSNVCLSN	60
Db	1	MAEAGLRGWLWALLRLLAQSEPYTTTHQPGYCAFYDECGRNPGLSGSLMTLSNVCLSN	60
Qy	61	TPARKITGDHLLILQKICPRLYTGNTQACCSAQOLVSEASLSITKALLTRCPACSNF	120
Db	61	TPARKITGDHLLILQKICPRLYTGNTQACCSAQOLVSEASLSITKALLTRCPACSNF	120
Qy	121	VNLHCHNTCSNPQSLFINVTRVAOLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180
Db	121	VNLHCHNTCSNPQSLFINVTRVAOLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180
Qy	181	ATLAVGTMCGVYGSAALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240
Db	181	ATLAVGTMCGVYGSAALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240
Qy	241	ARCNSQGGDVATCSQDCQCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
Db	241	ARCNSQGGDVATCSQDCQCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
Qy	301	LLVGFVRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Db	301	LLVGFVRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHPFFRTNQVILTAPNRSSY	420
Db	361	IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHPFFRTNQVILTAPNRSSY	420
Qy	421	RYDSLILGPKNFGSILDLILLLELQERLRLHQLVWSPEAQKNISLQDICYAPLNPDNT	480
Db	421	RYDSLILGPKNFGSILDLILLLELQERLRLHQLVWSPEAQKNISLQDICYAPLNPDNT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL	540
Qy	541	SCWADYGAPVFPFLAIGGYKDYSEAEALIMTFESLNNYPAGDPRIQAOKLWEAEFLEEM	600
Db	541	SCWADYGAPVFPFLAIGGYKDYSEAEALIMTFESLNNYPAGDPRIQAOKLWEAEFLEEM	600
Qy	601	RAFOREMAGMFOVTFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISALGSYSMSRV	660
Db			

Db 601 RAFQRMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGSSWSRV 560
Qy 661 MYDSKATLGGVAVVVLGAVMAAMGPFSSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MYDSKATLGGVAVVVLGAVMAAMGPFSSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Qy 721 YQRLPRPCEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTTALTSLAV 780
Db 721 YQRLPRPCEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTTALTSLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPOELPPPGQEGELLGFFQKAVAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPOELPPPGQEGELLGFFQKAVAPFL 840
Qy 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLFNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLFNRYFEVGAPV 900
Qy 901 YFVTTLYGNFSSEAGNNAICSSAGCNPFSTQKIQYATEFPQSVYLAIPASSWVDDFIDW 960
Db 901 YFVTTLYGNFSSEAGNNAICSSAGCNPFSTQKIQYATEFPQSVYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCCLKMSITMGSVRPSVEQPHKYLPLWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCCLKMSITMGSVRPSVEQPHKYLPLWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFVFPYTTNVFTEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFVFPYTTNVFTEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIVMLVDTVGFMAWDISYNAVSLNLVSAVGMSEVFSHTITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMLVDTVGFMAWDISYNAVSLNLVSAVGMSEVFSHTITRSPAISTKPTWLER 1200
Qy 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRNLNLTLLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRNLNLTLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVNPALALEOKRAEBAVAAMVASCNPHSPRVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPVNPALALEOKRAEBAVAAMVASCNPHSPRVSTADNIYVNHSPFGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 4
AAW79169
ID AAW79169 standard; protein; 1332 AA.
AC AAW79169;
DT 06-NOV-2001 (first entry)
DE Human protein SEQ ID NO 1831.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX Homo sapiens.
OS WO200157190-A2.
FN 09-AUG-2001.
PD 05-FEB-2001; 2001WO-US004098.
XX

XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX (HYSSE-) HYSEQ INC.
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX WPI: 2001-476283/51.
DR N-PSDB; AAK52302.
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX Claim 20; Page 4216-4219; 5221pp; English.
PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX SQ Sequence 1332 AA;
Query Match 99.8%; Score 6896; DB 4; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAPVDECKNPELSGLMTLSNVCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAPVDECKNPELSGLMTLSNVCLSN 60
Qy 61 TPARKITGDHLLQKICPRLYTGPTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLQKICPRLYTGPTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSFLINVTVAQOLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSFLINVTVAQOLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
Qy 181 ATLAVTGMCGVYGSALCNAQRLNFGDTEGNGLAPLDITFHLLEPQAVGSGIQLNEGV 240
Db 181 ATLAVTGMCGVYGSALCNAQRLNFGDTEGNGLAPLDITFHLLEPQAVGSGIQLNEGV 240
Qy 241 ARCNSQGDVATCSQDCAASCPAARPOALDSTFYLGQMPGSLVLIILCSFAVVTI 300
Db 241 ARCNSQGDVATCSQDCAASCPAARPOALDSTFYLGQMPGSLVLIILCSFAVVTI 300
Qy 301 LLVGRVAPARDKSKVNDPKGTSLSKLSSTHTLLGQPFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKVNDPKGTSLSKLSSTHTLLGQPFQCGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTLTDPVLEWSPNAPNSQAKSEKAFHQHFGFPFRFTQVILTAENRSY 420
Db 361 IPVVALAAGLVFTLTDPVLEWSPNAPNSQAKSEKAFHQHFGFPFRFTQVILTAENRSY 420
Qy 421 RYDSLLLGPKNPSGILDLLELLELQERLRLHLQVWSPFAQRNLSLQDICYAPLNPONT 480

Db 421 RYDSELLGPKFSGDLDLLELELEQLERLHQLVWSPQAQRNLSQDICYAPLNDPT 480
Qy 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTLAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTLAL 540
Qy 541 SCMDYGAVPFPFLAIGGYKGYKDYSEAEALIMTFSLNYPAGDPRPLAQAOKLWEAEFLBEM 600
Db 541 SCMDYGAVPFPFLAIGGYKGYKDYSEAEALIMTFSLNYPAGDPRPLAQAOKLWEAEFLBEM 600
Qy 601 RAFORMAGMFOVTTAERSLEDEINRTTADLPFAISYIVIFLYISLALGSSYSSWRV 660
Db 601 RAFORMAGMFOVTTAERSLEDEINRTTADLPFAISYIVIFLYISLALGSSYSSWRV 660
Qy 661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720
Qy 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAVRTPALTSGLAV 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAVRTPALTSGLAV 780
Qy 781 ILDFELLQMSAFVALLSLDSKQBEASRLDVCCCKQELPPQCGEGLLGFPQKAYAPFL 840
Db 781 ILDFELLQMSAFVALLSLDSKQBEASRLDVCCCKQELPPQCGEGLLGFPQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFNRYFVGVAPV 900
Db 841 LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFNRYFVGVAPV 900
Qy 901 YFVTTLGYNPSSEAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDPDFDW 960
Db 901 YFVTTLGYNPSSEAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDPDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCCLKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCCLKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAVSTSVNLTSQGVLASRFMAHKLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAVSTSVNLTSQGVLASRFMAHKLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPDTPAPEVFPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPDTPAPEVFPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIWMILVDTVGFMAWLDISYNAVSLNLSVAVGMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIWMILVDTVGFMAWLDISYNAVSLNLSVAVGMSVEFVSHITRSFAISTKPTWLER 1200
Qy 1201 AKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIPFRLMLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIPFRLMLLITLGLLHGLVFLPV 1260
Qy 1261 ILSVGDVNPALALEQKRAEEAAVAVMVASCPNHPRSVSTADNIYNHVSFEGSIKAGA 1320
Db 1261 ILSVGDVNPALALEQKRAEEAAVAVMVASCPNHPRSVSTADNIYNHVSFEGSIKAGA 1320
Qy 1321 ISNPLPNNGRQF 1332
Db 1321 ISNPLPNNGRQF 1332

RESULT 5
AAG65638
ID AAG65638 standard; protein; 1332 AA.
XX
AC AAG65638;
XX
DT 07-JAN-2002 (first entry)
XX
DE SSD-containing SSP1 protein sequence.

XX
KW SSD; sterol-sensing domain; human; liver; testis; brain; cancer;
KW nootropic; neuroprotective; antidiabetic; antiarteriosclerotic;
XX cytostatic; antilipemic; SSP1.
OS Homo sapiens.
XX
PN WO200170974-A1.
XX
PD 27-SEP-2001.
XX
PF 22-MAR-2001; 2001WO-JP002279.
XX
PR 24-MAR-2000; 2000JP-00088595.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Taniyama Y, Kita S, Komiyama T;
XX
XX WPI; 2001-611501/70.
XX N-PSDB; RAI66914.
XX
DR New sterol-sensing domain-containing protein for diagnosing and screening
PT candidate compounds in drug development for diabetes, obesity, cancer,
PT arteriosclerosis, hyperlipidemia and neurodegenerative disorders.
XX
PS Claim 7; Page 115-122; 171pp; Japanese.
XX
CC The invention provides a novel SSD (sterol-sensing domain)-containing
CC protein. The protein originates from human liver, human testis or human
CC brain. The protein can be expressed by standard recombinant methodology.
CC The proteins, encoded DNAs and antibodies are useful in diagnosis and
CC screening candidate compounds in drug development for diabetes, obesity,
CC cancer, arteriosclerosis, hyperlipemia, neurodegenerative disorders such
CC as Alzheimer's disease and neural disorders. The present sequence
CC represents the SSP1 protein which contains the SSD domain
XX
SQ Sequence 1332 AA;

Query Match 99.8%; Score 6896; DB 4; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Qy 61 TPARKITGDHLILLOKICPRLYTGNTOACCSAKOLVSLSEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGNTOACCSAKOLVSLSEASISITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAPQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAPQSYDSCSRVRPAA 180
Qy 181 ATLVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQLNEGV 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQLNEGV 240
Qy 241 ARCNEQGGDDVATCSCQDCAAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGGDDVATCSCQDCAAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTFELTDDPVELMSAPNSQARSEKAFHDQHFPGFPFTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTFELTDDPVELMSAPNSQARSEKAFHDQHFPGFPFTNQVILTAPNRSY 420
Qy 421 RYDSELLGPKNFGSGLDLDLLELELEQLERLHQLVWSPQAQRNLSQDICYAPLNDPT 480

Db 421 RYDLLLLGPKNPSGILDLDLLELLELQERLHLQVSPBAQRNLSLQDICVAPLNPONT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
Qy 541 SCMDYAGAPVPPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQWEEAFLEEM 600
Db 541 SCMDYAGAPVPPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQWEEAFLEEM 600
Qy 601 RAFORMAGMFOVTFATRSLEDEINRTTAEDLPATSYIVIFLYISIALGSSYSWSRV 660
Db 601 RAFORMAGMFOVTFATRSLEDEINRTTAEDLPATSYIVIFLYISIALGSSYSWSRV 660
Qy 661 MYDSKATLGLGGAVALGAVMAAGFFSYLGIRSSLVILQVPPFLVLSVGADNIFVYLE 720
Db 661 MYDSKATLGLGGAVALGAVMAAGFFSYLGIRSSLVILQVPPFLVLSVGADNIFVYLE 720
Qy 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTPALTSGLA 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTPALTSGLA 780
Qy 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVCKPQELPPPGQGGGLLGGFKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVCKPQELPPPGQGGGLLGGFKAYAPPL 840
Qy 841 LHWITRGVLLLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTLTGYNFSEAGMNAICSSAGCNPFSTQKIYATEPPEQSYLAIPASSWVDDFDW 960
Db 901 YFVTLTGYNFSEAGMNAICSSAGCNPFSTQKIYATEPPEQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNCHSITWGSVRPSVBEQHKYLPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNCHSITWGSVRPSVBEQHKYLPWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVLAASFMAHFKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLAASFMAHFKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAPFVPPYITNVFYEQVLTILPGLFMLSICLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAPFVPPYITNVFYEQVLTILPGLFMLSICLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSTVMLLVDTVGFMAWDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSTVMLLVDTVGFMAWDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200
Qy 1201 AKEATISMGSAVFAGVAMTNLPGIILVLGLAKAQILQIPFRLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVFAGVAMTNLPGIILVLGLAKAQILQIPFRLNLLITLGLLHGLVFLPV 1260
Qy 1261 ILSVGPDPVPALALEQKRAEEAAMVAVSCPNHPSRVSTADNIYNVHSPGSIKGAGA 1320
Db 1261 ILSVGPDPVPALALEQKRAEEAAMVAVSCPNHPSRVSTADNIYNVHSPGSIKGAGA 1320
Qy 1321 ISNPLPNNNGROF 1332
Db 1321 ISNPLPNNNGROF 1332

RESULT 6

AAM79168 standard; protein; 1359 AA.

XX AC AAM79168;

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 1830.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00663561.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI, 2001-476283/51.

XX N-PSDB; AAK52301.

Nucleic acids encoding polypeptides with cytokine-like activities, useful
in diagnosis and gene therapy.

XX Claim 20; Page 4213-4216; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
sequence listing were missing at the time of publication

XX SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLGLWLLALLRLAQSEPTTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60

Db 1 MAEAGLGLWLLALLRLAQSEPTTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60

Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTQACSAKQLVSLASLSITKALLTRCPACSDNF 120

Db 61 TPARKITGDHLLILLOKICPRLYTGPNTQACSAKQLVSLASLSITKALLTRCPACSDNF 120

Qy 121 VNLCHTCSNQSILFNTVTRVAQGLPAAVAYEAFYQHSFABQSYDCSRRVRPAA 180

Db 121 VNLCHTCSNQSILFNTVTRVAQGLPAAVAYEAFYQHSFABQSYDCSRRVRPAA 180

Qy 181 ATLAGVTGCVYGSALCNAQRWLNFGQDTGNGLAPLDTTHLLPQAVGSGIQPLNKG 240

Db 181 ATLAGVTGCVYGSALCNAQRWLNFGQDTGNGLAPLDTTHLLPQAVGSGIQPLNKG 240

Qy 241 ARCNEQDDVATCSCQDCCAASCPAIARPOALDSTFVLGQMPGSLVLIILCSVPFVVTI 300

Db	241	ARCNESQDDVATCSCQDCAASCPAIARPOLADSTFYLGQMPGSLVLIILICSVFAVVTI	300
Qy	301	LLVGFVRVAPARDKSMVDPKKGTSLSKLSFSHTLLGQFFQGMGTWVASWPLTILVLSV	360
Db	301	LLVGFVRVAPARDKSMVDPKKGTSLSKLSFSHTLLGQFFQGMGTWVASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFHDQHFQPPFRNQVILTAPNRSSY	420
Db	361	IPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFHDQHFQPPFRNQVILTAPNRSSY	420
Qy	421	RYDSLLGPKNPSGILDLDLLLELEQLERLRLHQLWSPQAQRNLSLQDICYAPLNPDNT	480
Db	421	RYDSLLGPKNPSGILDLDLLLELEQLERLRLHQLWSPQAQRNLSLQDICYAPLNPDNT	480
Qy	481	SLYDCCINSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGKTALAL	540
Db	481	SLYDCCINSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGKTALAL	540
Qy	541	SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWBEAFLEEM	600
Db	541	SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWBEAFLEEM	600
Qy	601	RAFQRMAGMFOVTPFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSSYSMSRV	660
Db	601	RAFQRMAGMFOVTPFAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSSYSMSRV	660
Qy	661	MYDSKATLGLGGAVALGAVNMAAGFFSYLGRSSLVILQVVPFLVLSVGADNFIPIFVLE	720
Db	661	MYDSKATLGLGGAVALGAVNMAAGFFSYLGRSSLVILQVVPFLVLSVGADNFIPIFVLE	720
Qy	721	YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFLGAITPMPAVRTFALTSGLA	780
Db	721	YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFLGAITPMPAVRTFALTSGLA	780
Qy	781	ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPQGGGELLGFPQKAYAPFL	840
Db	781	ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPQGGGELLGFPQKAYAPFL	840
Qy	841	LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFNRYFVVGAPV	900
Db	841	LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFNRYFVVGAPV	900
Qy	901	YFVTTLYGNFSSBAGMNAICSSAGCANNFSTQKIQVATEFPFQSYLAIPASSWVDVDFIDW	960
Db	901	YFVTTLYGNFSSBAGMNAICSSAGCANNFSTQKIQVATEFPFQSYLAIPASSWVDVDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDKCPSVTNSLNCIKNCMSITMGSVRPSVEQPHKYLIPWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDKCPSVTNSLNCIKNCMSITMGSVRPSVEQPHKYLIPWFLNDRP	1020
Qy	1021	NIKCPKGLAAYSTSVNLTSDGOVL-----ASRPMAYH 1053	
Db	1021	NIKCPKGLAAYSTSVNLTSDGOVLDTVALTSLPRLEYSGTISAHCNLYLLDSASRPMAYH 1080	
Qy	1054	KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITITNVFQEYLTILPEGL	1113
Db	1081	KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITITNVFQEYLTILPEGL	1140
Qy	1114	FMLSCLVPTFAVSCLLGLDRLSGLNLALSIVMILVDVTGFMALWDISVNAVSLNLNVS	1173
Db	1141	FMLSCLVPTFAVSCLLGLDRLSGLNLALSIVMILVDVTGFMALWDISVNAVSLNLNVS	1200
Qy	1174	AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1233
Db	1201	AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1260
Qy	1234	LIQIFFFRLLNLLITLLGLLHGLVLPVLVLSVGDVNPALALBQKRAEEAAVAVMVASCP	1293
Db	1261	LIQIFFFRLLNLLITLLGLLHGLVLPVLVLSVGDVNPALALBQKRAEEAAVAVMVASCP	1320
Qy	1294	NHPSRVSTADNIYVNHSPFSGIKGAGISNPLPNNGRQF 1332	

Db	1321	NHPSRVSTADNIYVNHSPFSGIKGAGISNPLPNNGRQF 1359	
RESULT 7			
ADJ27337			
ID	ADJ27337	standard; protein; 1359 AA.	
XX			
AC	ADJ27337;		
XX			
DT	20-MAY-2004	(first entry)	
XX			
DE	Human NPC1L1.		
XX			
KW	Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;		
KW	plasma membrane; transport signal; promoter;		
KW	sterol regulated element binding protein 1; SREBP1;		
KW	binding consensus sequence; transmembrane domain; sterol-sensing domain;		
KW	SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;		
KW	cholesterol absorption; serum cholesterol; hyperlipidaemia;		
KW	atherosclerosis; coronary heart disease; stroke; arteriosclerosis.		
XX			
OS	Homo sapiens.		
XX			
FN	WO2004009772-A2.		
PD	29-JAN-2004.		
XX			
PF	17-JUL-2003; 2003WO-US022467.		
XX			
PR	19-JUL-2002; 2002US-0397442P.		
XX			
PA	(SCHE) SCHERING CORP.		
XX			
PI	Altman SW, Murgolo NJ, Wang LQ, Graziano MP;		
DR	WPI; 2004-132945/13.		
DR	N-PSDB; ADJ27336.		
XX	GENBANK; AF192522.		
PT	New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting		
PT	or identifying antagonists of NPC1L1 for inhibiting intestinal		
PT	cholesterol absorption in a subject, or for treating elevated serum		
PT	cholesterol or stroke.		
XX			
PS	Example 7; SEQ ID NO 44; 125pp; English.		
XX			
CC	This sequence represents a Niemann-Pick disease, type C1, gene-like 1		
CC	(NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains		
CC	a motif ADJ27331 which acts as a trans-golgi network to plasma membrane		
CC	transport signal, and which exhibits limited tissue distribution and		
CC	gastrintestinal abundance. The human NPC1L1 promoter sequence contains a		
CC	sterol regulated element binding protein 1 (SREBP1) binding consensus		
CC	sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-		
CC	sensing domain (SSD) which is involved in sensing cholesterol levels,		
CC	possibly by a mechanism which involves direct cholesterol bonding. NPC1L1		
CC	has 42% amino acid sequence homology to human NPC1, a receptor		
CC	responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or		
CC	polynucleotides are useful for detecting or identifying antagonists of		
CC	NPC1L1, which can be used for inhibiting intestinal cholesterol		
CC	absorption in a subject, or for treating medical conditions including		
CC	elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary		
CC	heart disease, stroke or arteriosclerosis.		
XX			
SQ	Sequence 1359 AA;		

Query Match 99.5%; Score 6872.5; DB 8; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLUGLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN 60
Db 1 MAEAGLUGLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASLITKALLTRCPACSDNP 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASLITKALLTRCPACSDNP 120
QY 121 VNLHCHNTCSPNOSLFINTVTRVAQLGAGOLPAVVAEAFVQHSFASQSDYSCSRVRPAA 180
DB 121 VNLHCHNTCSPNOSLFINTVTRVAQLGAGOLPAVVAEAFVQHSFASQSDYSCSRVRPAA 180
QY 181 ATLAUVTGCVGYSALCNAQWLNFGQDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
DB 181 ATLAUVTGCVGYSALCNAQWLNFGQDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
QY 241 ARCNESQGDVATCSQDCAASCPATARPQALDSTFYLGQMPGSLVLLIILCSVFAVVTI 300
DB 241 ARCNESQGDVATCSQDCAASCPATARPQALDSTFYLGQMPGSLVLLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFPQGWGTVWASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFPQGWGTVWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPFPFRTNQVILTAPNRSY 420
QY 421 RYDSLLGPNKNSGILDLDDLELLELQRLRLHQLWSPEAQRNLSLQDICYAPLNPDT 480
DB 421 RYDSLLGPNKNSGILDLDDLELLELQRLRLHQLWSPEAQRNLSLQDICYAPLNPDT 480
QY 481 SYDDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDHFYLCANAPLTKQGTALAL 540
DB 481 SYDDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDHFYLCANAPLTKQGTALAL 540
QY 541 SCHADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
DB 541 SCHADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAFORMAGMFWTFAERSLDEINRTTAEDLPATSYIIVIFLYISALGYSYSWSRV 660
DB 601 RAFORMAGMFWTFAERSLDEINRTTAEDLPATSYIIVIFLYISALGYSYSWSRV 660
QY 661 MVDKATILGGVAVVILGAVMAAMGFFSYLGRSSILVILQVVPFLVSVGADNIFIVLE 720
DB 661 MVDKATILGGVAVVILGAVMAAMGFFSYLGRSSILVILQVVPFLVSVGADNIFIVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEALICFELGALTMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEALICFELGALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCKPQELPPQGGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCKPQELPPQGGGLLGFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYSMCHISVGLQDELALPKDSYLLDYFLFLNRYFVGAPV 900
DB 841 LHWITRGVLLFLALFGVLSYSMCHISVGLQDELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTLGYNFSBAGMNAICSSAGCNFSTOKIOVATEBPPQSYLAIIPASSWVDPDFIDW 960
DB 901 YFVTTLGYNFSBAGMNAICSSAGCNFSTOKIOVATEBPPQSYLAIIPASSWVDPDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMCSITMGSVRPSVEQFKHYPFLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMCSITMGSVRPSVEQFKHYPFLWFLNDRP 1020
QY 1021 NTKCPKGGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
DB 1021 NTKCPKGGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEVFPYTTITNVFVEQVLTILPEGL 1113
DB 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEVFPYTTITNVFVEQVLTILPEGL 1113
QY 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEVFPYTTITNVFVEQVLTILPEGL 1140
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEVFPYTTITNVFVEQVLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRSGLLNLISIVMLVDTVGFPMALWDISYNAVSLINLVS 1173

DB 1141 FMLSCLVPTFAVSCLLGLDLRSGLLNLISIVMLVDTVGFPMALWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATTISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATTISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIPFFRLNLITLGLLHGLVFLPVILSYVYVDPVNPALAEQKRAEBAVAAMVVASCP 1293
DB 1261 LIQIPFFRLNLITLGLLHGLVFLPVILSYVYVDPVNPALAEQKRAEBAVAAMVVASCP 1320
QY 1294 NHPSRVSTADNTIYNHVSFEGSIKGAGAINPFLPNNGRQF 1332
DB 1321 NHPSRVSTADNTIYNHVSFEGSIKGAGAINPFLPNNGRQF 1359
RESULT 8
ADY60780
ID ADY60780 standard; protein; 1359 AA.
XX AC ADY60780;
XX AC ADY60780;
DT 19-MAY-2005 (first entry)
XX DE Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 44.
XX KW Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX OS Homo sapiens.
XX PN WO2005015988-A1.
XX PD 24-FEB-2005.
XX PF 16-DEC-2003; 2003WO-US040113.
XX PR 17-JUL-2003; 2003US-00621758.
PR 22-AUG-2003; 2003US-00646301.
PR 16-SEP-2003; 2003US-00663208.
XX (SCHE) SCHERING CORP.
XX PI Altman SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2005-284403/29.
XX N-PSDB; ADY60779.
PT Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
PS Disclosure; SEQ ID NO 44; 146pp; English.
XX The invention relates to an isolated Niemann-Pick C1-like protein 1
CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
CC amino acid sequence appearing as ADY60738,ADY60740 or ADY60748 (Rat,
CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
CC encoding NPC1L1 (appearing as ADY60737,ADY60739 or ADY60747), a
CC recombinant vector comprising the polynucleotide, a host cell comprising
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
CC produce any functional NPC1L1 protein), an offspring or progeny of the
CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
CC dosage form, and information indicating that NPC1L1 is a target of

XX Identifying ligand of NPC1l1 for stimulating the activity of NPC1l1, by
PT contacting NPC1l1 with detectably labeled substituted 2-azetidinone
PT glucuronide and a candidate compound and determining if compound binds to
PT human NPC1l1.

XX Disclosure; SEQ ID NO 44; 215pp; English.

XX The invention relates to identifying a ligand of NPC1l1. The method
CC involves contacting human NPC1l1 with a detectably labeled substituted 2-
CC azetidinone glucuronide and a candidate compound and determining whether
CC the candidate compound binds to human NPC1l1. In identifying a ligand of
CC NPC1l1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10
CC nM or lower. The detectably labeled substituted 2-azetidinone glucuronide
CC is 35 S-labeled compound 2, given in the specification. NPC1l1 ligands
CC are useful for stimulating or blocking the activity of NPC1l1, and for
CC treating conditions caused or mediated by NPC1l1. It is useful for
CC reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart
CC disease, stroke, or arteriosclerosis. The present sequence represents the
CC amino acid sequence of human NPC1l1 as disclosed under GenBank accession
CC number AF192522.

XX Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 9; Length 1359;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY	1	MAEAGLRGWLWALLRLAQAQSPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN	60
DB	1	MAEAGLRGWLWALLRLAQAQSPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN	60
QY	61	TPARKITGDHLILLOKICPRLTYGTNPNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNF	120
DB	61	TPARKITGDHLILLOKICPRLTYGTNPNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNF	120
QY	121	VNLHCHNTCSNPQSLFINVTRVAQAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA	180
DB	121	VNLHCHNTCSNPQSLFINVTRVAQAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA	180
QY	181	ATLVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDTFHLLERQAVGSGIQPLNEGV	240
DB	181	ATLVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDTFHLLERQAVGSGIQPLNEGV	240
QY	241	ARCNESQGDVATCSQDCAACSPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
DB	241	ARCNESQGDVATCSQDCAACSPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
QY	301	LLVGRVPAPARDKSRKMDPKKTSLSDKLSFSTHTLLGQFPQGWGTTWASWPLTILVLSV	360
DB	301	LLVGRVPAPARDKSRKMDPKKTSLSDKLSFSTHTLLGQFPQGWGTTWASWPLTILVLSV	360
QY	361	IPVVALAAGLVFTELTDPVELWSPNSQARSEKAFPHQHPFPRTNQVILTAPNRSY	420
DB	361	IPVVALAAGLVFTELTDPVELWSPNSQARSEKAFPHQHPFPRTNQVILTAPNRSY	420
QY	421	RYDLSLLGPKNFGSLDLDLLELELELRLHLQWSPQAORNISLQDICVAPLNPNPT	480
DB	421	RYDLSLLGPKNFGSLDLDLLELELELRLHLQWSPQAORNISLQDICVAPLNPNPT	480
QY	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKMDHFLYCANAPLTFKDGTTALAL	540
DB	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKMDHFLYCANAPLTFKDGTTALAL	540
QY	541	SCMADYGAPVFPFLAIGGKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM	600
DB	541	SCMADYGAPVFPFLAIGGKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM	600
QY	601	RAFORRMAGMFOVTTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSYSWSRV	660
DB	601	RAFORRMAGMFOVTTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSYSWSRV	660
QY	661	MYDSKATILGLGGAIVVLGAVMAAMGFFSYLGRSSSLVILQVVPFLVLSVGADNIFIFVLE	720

DB	661	MYDSKATILGLGGAIVVLGAVMAAMGFFSYLGRSSSLVILQVVPFLVLSVGADNIFIFVLE	720
QY	721	YORLPRRPGEPREVEHIGRALGRVAPSMLLCSLSEACPFPLGALTMPAPVTRTALTSLAV	780
DB	721	YORLPRRPGEPREVEHIGRALGRVAPSMLLCSLSEACPFPLGALTMPAPVTRTALTSLAV	780
QY	781	ILDPLQLMSAFVALLSLDSKROEASRLDVCCCKPQELPPPPQGGEGLLGFFQKAYAPFL	840
DB	781	ILDPLQLMSAFVALLSLDSKROEASRLDVCCCKPQELPPPPQGGEGLLGFFQKAYAPFL	840
QY	841	LHWITRGVWLLFLALFGVLSYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAVP	900
DB	841	LHWITRGVWLLFLALFGVLSYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAVP	900
QY	901	YVVTTLGYNFSSEAGNNAICSSAGCWNFSFTOKIYATFEPPQSYLAIPASSWVDDPDM	960
DB	901	YVVTTLGYNFSSEAGNNAICSSAGCWNFSFTOKIYATFEPPQSYLAIPASSWVDDPDM	960
QY	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLFWFLNDRP	1020
DB	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLFWFLNDRP	1020
QY	1021	NIKCPKGLAAYSTSVNLTSDQVLT-----ASRPMATH	1053
DB	1021	NIKCPKGLAAYSTSVNLTSDQVLT-----ASRPMATH	1053
QY	1054	KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL	1113
DB	1054	KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL	1113
QY	1114	FMLSCLVPTFAVSCLLGLDLRLSGLLMLLSIVMLVDTVGFMAWLDISYNAVSLINVS	1173
DB	1114	FMLSCLVPTFAVSCLLGLDLRLSGLLMLLSIVMLVDTVGFMAWLDISYNAVSLINVS	1173
QY	1174	AVGMSVEFVSHITRSFAISTKPTWLERAKBATISMGSAVFAGVAMTNLPGLILVLAKAQ	1233
DB	1174	AVGMSVEFVSHITRSFAISTKPTWLERAKBATISMGSAVFAGVAMTNLPGLILVLAKAQ	1233
QY	1234	LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVYGVDPNPALEOKRAEBAVAAMVVASCP	1293
DB	1234	LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVYGVDPNPALEOKRAEBAVAAMVVASCP	1293
QY	1294	NHPSRVSTADNIIYVNHSPGSIKGAISNFIPLNNGRQF	1332
DB	1294	NHPSRVSTADNIIYVNHSPGSIKGAISNFIPLNNGRQF	1332
QY	1332	NHPSRVSTADNIIYVNHSPGSIKGAISNFIPLNNGRQF	1359
DB	1332	NHPSRVSTADNIIYVNHSPGSIKGAISNFIPLNNGRQF	1359

RESULT 10
ABG22693
ID ABG22693 standard; protein; 1344 AA.
XX AC ABG22693;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #22684.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX OS Homo sapiens.
XX WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.

Qy	1079	DLRKVPCTDPAPFVFPYTTINVFYEQYLITLPBGLFMLSLCLVPTFAVSCLLGLDLSRG	1138
Dd	1080q	DLRKVPCTDNFEFPYTTISNVFYQQYLTVLPGIFTLALCFPFTFVCVLLGLDMCSG	1139
Qy	1139	LNLNLSVTMLVDTVGPMAWDLSYNVAVSLINLVSAGMSVEFVSHITRSFALSTKPTWL	1198
Dd	1140	: :	1199
Qy	1199	ERAKEATISGSAVFAVAMTNIPGILVLGLAKAQLIQIFFPRNLNLTLLGLHLGVFL	1258
Dd	1200	ERAKDAYTFMGSAVFAVAMTNPEGILGLFAOQLTIQIFFRNLNLTLLGLHLGVFL	1259
Qy	1259	PVILSYGGPDVNPALALEQKRAREAAVMVASCPNPHSRVSTADNIYNNHSPGST-KG	1317
Dd	1260	PVILSYLGPDVNQALVCEEKLASEA-AVAPESPQPYPSPADAN--VNYGAPELAHG	1316
Qy	1318	AGAISNPLPNNGROF	1332
Dd	1317	ANAARSLSPKSDOKF	1331

RESULT 13

RESOL 13
AEB93568
ID AEB93568 standard; protein; 1331 AA.

AC AEB93568;

06-OCT-2005: (first entry)

Rat NPC111 protein.

XX Protein engineering; NPC11; antidiabetic; antiarteriosclerotic;
KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
KW arteriosclerosis; rat.

OS *Rattus* sp.

WO2005069900-A2.

04-AUG-2005.

AA
PF 14-JAN-2005; 2005WO-US001469.

16-JAN-2004: 2004US-0537341P.

PA (MERI) MERCK & CO INC.

XX PI Garcia-Calvo M:

XX
DR WPI: 2005-564070/57.

DR N-PSDB; AEB93567.
XX

identifying ligand of NPC111 for stimulating the activity of NPC111, by contacting NPC111 with detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining if compound binds to human NPC111.

Example 2: SEO ID NO 2: 215pp: English.

The invention relates to identifying a ligand of NPC1L1. The method involves contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining whether the candidate compound binds to human NPC1L1. In identifying a ligand of NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10 nM or lower. The detectably labeled substituted 2-azetidinone glucuronide is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands are useful for stimulating or blocking the activity of NPC1L1, and for treating conditions caused or mediated by NPC1L1. It is useful for reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart disease, stroke, or arteriosclerosis. The present sequence represents a rat NPC1L1, a N-glycosylated protein.

SQ Sequence 1331 AA;

QY 1019 RPNKCPKGGGLAAYSTSNLSDGQVLASRFMAYHKLKNSQDYTEALRAARELAANITA 1078
DB 1020 PPNIRCPKGGGLAAYRTSNLSSDQVIAQPMAYHKLPLRNSQDFTALRASRLAANITA 1079
QY 1079 DLKVPKGTDPAFEPVPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
DB 1080 DLKVPKGTDPNPEVFPYTIISNVFYQOYLTVLPEGIFTALCFVPTFVVCYLLGLDMCSG 1139
QY 1139 LNLLSIWMILVDVTGFMALWDISNAVSLNLVSAVGMSEVFSVSHITRSPAIKTKPTWL 1198
DB 1140 ILNLSIIMILVDITGLMAVGISNAVSLNLVTAVGMSVEFVSHITRSPAVSTKPTRL 1199
QY 1199 ERAKATISMSAGVAGVAMTNLPGLVLGLAKAQLIQIFPRLNLLITLGLHLGLVFL 1258
DB 1200 ERAKATVFMGSAVAGVAMTNFPGILITLGAQALIQIFPRLNLLITLGLHLGLVFL 1259
QY 1259 PVILSYVGPDPNPAALQKRAEAAVAVMVASCPNHPRSRVSTADNIVVNSHSPFGSI-KG 1317
DB 1260 PVVLSYLGPDVNQALVQBEKLAESA-AVAPBPCPOYSPADADAN--VNYGFAPELAH 1316
QY 1318 AGAISNFPNNGROF 1332
DB 1317 ANAARSSLPKSDQKF 1331

RESULT 14
ID ADJ27305
XX ADJ27305 standard; protein; 1333 AA.
AC ADJ27305;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse NPC1L1.
XX
KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.
XX
OS Mus sp.
XX
PN WO2004009772-A2.
XX
PD 29-JAN-2004.
XX
PF 17-JUL-2003; 2003WO-US022467.
XX
PR 19-JUL-2002; 2002US-0397442P.
XX
PA (SCHE) SCHERING CORP.
XX
PI Altman SW, Murgolo NJ, Wang LQ, Graziano MP;
XX
DR WPI; 2004-132945/13.
DR N-PSDB; ADJ27304.
XX
PT New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting
PT or identifying antagonists of NPC1L1 for inhibiting intestinal
PT cholesterol absorption in a subject, or for treating elevated serum
PT , cholesterol or stroke.
XX
PS Claim 1; SEQ ID NO 12; 125pp; English.
XX
CC This sequence represents a Niemann-Pick disease, type C1, gene-like 1
CC (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains
CC a motif ADJ27331 which acts as a trans-golgi network to plasma membrane
CC transport signal, and which exhibits limited tissue distribution and
CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a

CC sterol regulated element binding protein 1 (SREBP1) binding consensus
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
CC sensing domain (SSD) which is involved in sensing cholesterol levels.
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1
CC has 42% amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.
XX
SQ Sequence 1333 AA;

Query Match 78.3%; Score 5407; DB 8; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;
QY 4 AGLRGLLWALLRLAQSEPTTTHQPGYCAFVDECKNPGLSMTLSNVSCLSNTPA 63
DB 3 AAWGMLLWALLNSAQGLYTPTHKAGPCTFYEECKNPGLSMTLSNVSCLSNTPA 62
QY 64 RKITGDHLILQKICPLRYTGNP-TQACCSAKQLVLSLEASLITKALLTRCPACSDNPNV 122
DB 63 RHVTGDHLALLQRCVPRLYNGPNDTYACCSTKQLVSLDSSLITKALLTRCPACSENFVS 122
QY 123 LHCHTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFVQHSFASQSDSCSRVPPAAAT 182
DB 123 IHCHTCSPPQSLFINVTRVQRPQQLPAVVAEAFYQSFABKAVESCSRVRIPAAAS 182
QY 183 LAVGTMGVYGSALCNAQRMLNFQDGTGNGLAPLDTITPHLEPQAVGSGIQPLNEGVAR 242
DB 183 LAVGSMGVYGSALCNAQRMLNFQDGTGNGLAPLDTITPHLEPQALADGKMLDGKITP 242
QY 243 CNESQGDVATCSQDCQCAASCPAIPALQALDSTFVLCMPGSIIVLIIILCSVFAVVTILL 302
DB 243 CNESQGEDSAACQDCQCAASCPVPPPPALRSPFYMGRMPGMLALIIIFTAVFVLLSVVL 302
QY 303 VGFVAPARDKSKMVDPKKGTSLSDKLSFTHTLLGQFQGWGTWVASWPLTILVLSVIP 362
DB 303 VTLRVASNRNKKTAGSQEAENLPKRFRFSPHTVLGRFFESWGTFVASWPLTILVLSFIV 362
QY 363 VVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQFPFRNTQVILTPNRSSTRY 422
DB 363 VIALSVGLTFIELTTDPVELMSAPKSQARSEKAFHDEHFQFPFRNTQVILTPNRSSTRY 422
QY 423 DSSLGPKNFGSILDLILLLELLELQERLRLHQLVQWSPEAQRNISIQDICYAPLNPDNTSL 482
DB 423 DSSLGPKNFGSILDLILLLELLELQERLRLHQLVQWSPEAQRNISIQDICYAPLNPDNTSL 482
QY 483 YDCCINSLLQVFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTLALSC 542
DB 483 TDCCVNSLLQVFQNNHTLLLTANQTLNGQTSQVDWKDHFLYCANAPLTFKDGTLALSC 542
QY 543 MADYGAPVFPFLAIGGYKGOYSEBALIMTFSLNNYPAGDPLRQAOKLWEAEPLLEMPRA 602
DB 543 IADYGAPVFPFLAVGGYQTDYSEAEALITITSINNYPADDDPRMAHAKLWEAEPLKEMOS 602
QY 603 FORRMAGMFOVTTFAERSLEDEINTEADLPFATSYIVIFLYISLALGSYSSSRVMV 662
DB 603 FORSTADKFPQIAFSAERSLEDEINTEADLPFATSYIVIFLYISLALGSYSSSRVMV 662
QY 663 DSKATILGLGGVAVLGVAMMAAGFFSYLGRSSVLQVVPFLVLSVGADNIFIFVLEYQ 722
DB 663 DSKATILGLGGVAVLGVAMMAAGFFSYLGRSSVLQVVPFLVLSVGADNIFIFVLEYQ 722
QY 723 RLPRRPGEPREVIHIGRALGRVAPSMMLCSLSAICFFLIGALTMPAVRTFALTSGVLVL 782
DB 723 RLPRMPGEQREAHIGRTLGSVAPSMMLCSLSAICFFLIGALTMPAVRTFALTSGVLVL 782
QY 783 DFLQWMSAVALLSLDLSKQKQASRLDVCCKVPQELPPPGQEGGLLGFQKAYAPFLHLH 842
DB 783 DFLQWTAFAVALLSLDLSKQKQASRPDVCCKPSRNLPPPKQKQEGGLLCCFFRKYITPFLHLH 842


```
Db 423 DLSLLGPKNFSGILSDLLDQLLELLEQERHLQWMSHAQRNISLQDICYAPLPHNTSL 482
Qy 483 YDCCNLSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGDTALALSC 542
Db 483 TDCCVNSLQYFQNNHTLLLTANQTLNGQTSQVDWKDHFLYCANAPLTYKDGDTALALSC 542
Qy 543 MADYGAPVPPFLAIGYKGYSEBALIMFSLNNYPAGDPRLAQAQKWEAEFLLENRA 602
Db 543 IADYGAPVPPFLAVGQYQGTYSAEALITFSINNYPADDPMAHAKLWEAEFLKEMQS 602
Qy 603 FORMMAGHFOVTFTEAERSLEDEINTTAEEDLPIFATSVIVFLYISLALGSYSSMSRVYV 662
Db 603 FORSTADKFKIAPSAERSLEDEINTTQDLFPFAISLIVFLYISLALGSYSSRSRVAV 662
Qy 663 DSKATILGGGVAVLGVMAAMGFFSYLGIRSLVILQVVPFLVSLVGADNIFIFVLSYQ 722
Db 663 DSKATILGGGVAVLGVMAAMGFYSLGVPSLVIQVVPFLVLAAGADNIFIFVLSYQ 722
Qy 723 RLPRPGEPREVIHGRALGRVAPSMMLCSLSAICFFLIGALTPMPAVRTFALTSLGLVIL 782
Db 723 RLPRMPGEQREAHGRTLGSVAPSMMLCSLSAICFFLIGALTSMPAVRTFALTSLGLIF 782
Qy 783 DFLLOMSAFVALLSLDSKQREASRLDVCCVCPQBLPPPGQEGLLGLGFFQKAYAPFLH 842
Db 783 DFLQWTAFAVALLSLDSKQREASRDVCCFSRNLPPPKQKEGLLCCFFRKYITPFLH 842
Qy 843 WITRGVLLFLALFGVSLYSWCHISVGLDQELALPKOSYLDYDFLFLNRYFEVQAPYV 902
Db 843 RFRPVLVLLFLVLFGANLYLNCNISVGLDQDLALPKOSYLDYDFLFLNRYLEVGPVYV 902
Qy 903 VTLGLNFSSEAGMNAICSSAGCNFSPQKIQAYTEPEQSYLAIPASSWDDFDWLT 962
Db 903 DTTSGYNFTEAGMNAICSSAGCESFSLTKIQYASEFPNQSYVAIAASSWDDFDWLT 962
Qy 963 P-SSCCRLYISGPNKDKPCPTVNSLNCNKMSITMGSVRPSVEQFHKYLFPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDEFCPTDTSFNCLKNCNRTLGPVPTTEQFHKYLFPWFLNDTPN 1022
Qy 1022 IKCPKGLAAYSTSVNLTDGQVLASRFMAHYHKLKNSQDYTEALRAARELANITADUR 1081
Db 1023 IRCEPKGLAAYRTSVNLSDDQIIASQFMAYHKPLRNSQDFTALRAASRLAANITAEUR 1082
Qy 1082 KVPDTPAPEVPPYTTINVPYEQYTLILPGLFMLSCLVPTFAVSCILGLDLRSLAN 1141
Db 1083 KVPDTPNPEVPPYTTISNVPYQYQYLVLPBEGIFTALCFVPTPVVCYLLGLDIRSGILN 1142
Qy 1142 LLSIVMLVDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1201
Db 1143 LLSIIMLVDTIGLMAVWGISYNAVSLINLVTAGVMSVEFVSHITRSFAVSTKPTRLER 1202
Qy 1202 KEATISGSAVPAVAVMTNLPGLILVGLAKAQILQIFPFLNLLITLGLLHGLVLPVY 1261
Db 1203 KDATIFWGSVAVFAGVAMTNPGLIILGFAQAQILQIFPFLNLLITLGLLHGLVLPVY 1262
Qy 1262 LSVVGDVNPALAEOKRAEEAAVAVWVACPNHPSRVSTADNIYNHSPFEGS- IKGAGA 1320
Db 1263 LSVLGDVNPQALVLEKATEA-AMVSEPCQYPPFADANTSDYVNGENPEFIFEINA 1321
Qy 1321 ISNFLPNNGRQF 1332
Db 1322 ASSSLPKSDQKF 1333
```

RESULT 16
AEB93578
ID AEB93578 standard; protein; 1333 AA.

XX AEB93578;

AC AEB93578;

XX 06-OCT-2005 (first entry)

DT 06-OCT-2005 (first entry)

XX 06-OCT-2005 (first entry)

DE Mouse NPC1L1 protein.

```
XX Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;  
KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;  
KW arteriosclerosis; mouse.  
XX Mus sp.  
XX WO2005069900-A2.  
XX 04-AUG-2005.  
XX 14-JAN-2005; 2005WO-US001469.  
XX 16-JAN-2004; 2004US-0537341P.  
XX (MERI ) MERCK & CO INC.  
XX Garcia-Calvo M;  
XX MPI; 2005-564070/57.  
XX N-PSDB; AEB93577.  
XX Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by  
PT contacting NPC1L1 with detectably labeled substituted 2-azetidinone  
PT glucuronide and a candidate compound and determining if compound binds to  
PT human NPC1L1.  
XX Example 2; SEQ ID NO 12; 215pp; English.  
XX The invention relates to identifying a ligand of NPC1L1. The method  
CC involves contacting human NPC1L1 with a detectably labeled substituted 2-  
CC azetidinone glucuronide and a candidate compound and determining whether  
CC the candidate compound binds to human NPC1L1. In identifying a ligand of  
CC NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10  
CC nM or lower. The detectably labeled substituted 2-azetidinone glucuronide  
CC is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands  
CC are useful for stimulating or blocking the activity of NPC1L1, and for  
CC treating conditions caused or mediated by NPC1L1. It is useful for  
CC reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart  
CC disease, stroke, or arteriosclerosis. The present sequence represents a  
CC mouse NPC1L1, a N-glycosylated protein.  
XX Sequence 1333 AA;  
Query Match 78.3%; Score 5407; DB 9; Length 1333;  
Best Local Similarity 77.2%; Pred. No. 0;  
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;  
Qy 4 AGLRGWLLWALLRLLAQSEPYTHIQPGYCAPFYDECGKNPELSGSLMTLSNVSCLSNTPA 63  
Db 3 AAWQWLLWALLLNSAQGELYTPTHKAGFCTPYEECGKNPELSGGLTSLSNISCLSNTPA 62  
Qy 64 RKTGDHLLILLOKICPRLYTGNN-TOACCSAKQLVSLASLSITKALLTRCPACSDNPNV 122  
Db 63 RHVTGDHLLALQVRCFRLYNGPNDYACCTKQLVSLDSSLITKALLTRCPACSENFVS 122  
Qy 123 LHCHNTCSNPQSLFNTVTRVAQLGAGQLPVAVYEAFAHQHSAEQSYDCSRVRVPAAT 182  
Db 123 IHCHNTCSPDQSLFINVTRVQDFQQLPAVVAEAFYQSFPAEKAYESCSRVIPAAAS 182  
Qy 183 LAVGTWCGYVGSALCNAQRWLNFGDGTGNGLAPLDTITFHLLPFGQAVSGIQPLNEGVAR 242  
Db 183 LAVGSMCGYVGSALCNAQRWLNFGDGTGNGLAPLDTITFHLLPFGQALADGMKPLDOKITP 242  
Qy 243 CNESOGDDVATCSCODCAASCPAIARPOALDSTFYLGOMPGSGSLVLIILCSFVAVVTILL 302  
Db 243 CNESOGEDSAACSCQDCAASCPIVPPPPALRPSFYMGPMFWLALIIITAVFVLSVVL 302  
Qy 303 VGFVRVAPARDKSMVDPKKGTSLSKLSPTHTLLGQFPGQGTWVASWPLTILVLSVIP 362  
Db 303 VYLRVASNENKNTAGSQEAPNLPKRKRPSPTVLGRPFESWGTWRVASWPLTVLALSFI 362  
Qy 363 VVALAAGLVPTTELTDTDPVELMSAPNSQARSEKAFHQHGFPPRTNQVILTAPNRSYRY 422
```


Db 363 VIALSVGLTFFIELTTPVELMSAPSKQARKEKAFDEHFGPPFRNQIFVTAKNRSSYKY 422
Qy 423 DSLILGPKNFSGILDLILLELELLELQERHLQVWSPQAQRNLSQDICYAPLNDSLSL 482
Db 423 DSLILGPKNFSGILSDLLLOELLELQERHLQVWSPQAQRNLSQDICYAPLNDSLSL 482
Qy 483 YDCINSLLQYQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTTALALSC 542
Db 483 TDCCVNSLLQYQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTTALALSC 542
Qy 543 MADYCAPVFPFLAIGYKGYSEABALIMFSLANNYPAGDRLAQAKLWEAFLEEMRA 602
Db 543 IADYCAPVFPFLAVGGYQGTDYSEBALIITFSINYPADDDPRMAHAKLWEAFLEEMQOS 602
Qy 603 FORMMAGMFQVFTTAERSLEDEINRTAEDLPIFATSVIVIFLYISLALGSVSSSRVWV 662
Db 603 FORSTADKFIQAFSAERSLEDEINRTIQLDPFAISLIVLYISLALGSVSRSRVAV 662
Qy 663 DSKATLGLGGVAVVLGAVVAMGFYSLGVPSLVIIQVVPFLVAVGADNIFIFVLEVQ 722
Db 663 DSKATLGLGGVAVVLGAVVAMGFYSLGVPSLVIIQVVPFLVAVGADNIFIFVLEVQ 722
Qy 723 RLPRPGEPREVHIGRAIGRVAPSMILCSLSEACFFFLGALTPMPAVRTFALTSGIIVL 782
Db 723 RLPRPGEPREVHIGRAIGRVAPSMILCSLSEACFFFLGALTPMPAVRTFALTSGIIVL 782
Qy 783 DFLQMSAFVALLSDSKQREASRLDCCVCKPQBLPPGQEGELGLGFFOKAYAPFLH 842
Db 783 DFLQMSAFVALLSDSKQREASRLDCCVCKPQBLPPGQEGELGLGFFOKAYAPFLH 842
Qy 843 WITRGVLLFLALFGVSLYSMCHISVGLDQELAPKDSYLDYFLFLNRYFEVGPVYF 902
Db 843 RPIRPVLLFLALFGVSLYSMCHISVGLDQELAPKDSYLDYFLFLNRYFEVGPVYF 902
Qy 903 VTLIGYNSSBAGNAICSSACGNFSTOKIYATEPQSYLAIPASSWVDDIDWLT 962
Db 903 DTTSGYNSTBAGNAICSSACGNFSTOKIYATEPQSYLAIPASSWVDDIDWLT 962
Qy 963 P-SSCCRLVIGSPNCKPCPTVNSLNCNKMSITMGSVRPSVBOFHKLFWFLNDRPN 1021
Db 963 PSSCCRLVIGSPNCKPCPTVNSLNCNKMSITMGSVRPSVBOFHKLFWFLNDRPN 1022
Qy 1022 IKCPKGLAAVSTSVNLSDGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLR 1081
Db 1023 IRCPKGLAAVSTSVNLSDGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLR 1082
Qy 1082 KVPGTDPAFVFPYITITVFEQYITLPEGLFMLSCLVPTTFAVSCILLGLDLRSGLIN 1141
Db 1083 KVPGTDPAFVFPYITITVFEQYITLPEGLFMLSCLVPTTFAVSCILLGLDLRSGLIN 1142
Qy 1142 LLSIIMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPASTKPTLERRA 1201
Db 1143 LLSIIMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPASTKPTLERRA 1202
Qy 1202 KEATISMGSAVFAGVAMTFLGILVLGLAKAQLIQIFFFRLNLIITLGLLHGLVFLPVI 1261
Db 1203 KDATIFMGSAVFAGVAMTFLGILVLGLAKAQLIQIFFFRLNLIITLGLLHGLVFLPVI 1262
Qy 1262 LSYGPDVNPALALBQKABEAAVAMVASCNPHSPRSTADNIVYNSHFRGS-IKGAGA 1320
Db 1263 LSYGPDVNPALALBQKABEAAVAMVASCNPHSPRSTADNIVYNSHFRGS-IKGAGA 1321
Qy 1321 LSNLPNNRGQF 1332
Db 1322 ASSSLPKSDQKF 1333

RESULT 17
ABG22691
ID ABG22691 standard; protein; 982 AA.
XX
AC ABG22691;

XX 18-FEB-2002 (first entry)
DT Novel human diagnostic protein #22682.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
DE food supplement; medical imaging; diagnostic; genetic disorder.
KW Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS86878.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 53050; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 982 AA;

Query Match 64.6%; Score 4466; DB 4; Length 982;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 867; Conservative 1; Mismatches 9; Indels 8; Gaps 3;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGNPELSGLMTLSNVCSLSN 60
Qy 61 TPARKITGDHLLILQKICPRLYTGPNTQACCSAQQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGPNTQACCSAQQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNQSILFINVTRVAQIGAGQLPAVVAYEAFYQHSFAEQSDSCSRVRPAA 180
Db 121 VNLHCHTCSNQSILFINVTRVAQIGAGQLPAVVAYEAFYQHSFAEQSDSCSRVRPAA 180

Qy	181	ATLAVGTCMGVYGALCNQAWRLNFQDGTGNGLAPLDITTHLLEPGQAVSGIQLPNEG	240
Db	181	ATLAVGTCMGVYGALCNQAWRLNFQDGTGNGLAPLDITTHLLEPGQAVSGIQLPNEG	240
Qy	241	ARCHESQDDVATCSQDCASCP--AIARPAQDSTFYLGOMPGSLVLIILCSVFVV	298
Db	241	ARCHESQDDVATCSQDCASCP--AIARPAQDSTFYLGOMPGSLVLIILCSVFVV	298
Qy	299	TILLVGRFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWVASMPLTILVL	358
Db	299	TILLVGRFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWVASMPLTILVL	358
Qy	359	SVIPVVALAAGLVFTELTTPVELWSAPNSQARSEKAFHDQHFQFPFRFTQVILTAPNRS	418
Db	359	SVIPVVALAAGLVFTELTTPVELWSAPNSQARSEKAFHDQHFQFPFRFTQVILTAPNRS	418
Qy	419	SYRYSLLGPKNSGIILDLLELELQERLRLHQLWSPQARNISLQDICYAPLNPD	478
Db	419	SYRYSLLGPKNSGIILDLLELELQERLRLHQLWSPQARNISLQDICYAPLNPD	478
Qy	479	NTSLYDCCINSLLQYFQNNRTILLITANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTAL	538
Db	479	NTSLYDCCINSLLQYFQNNRTILLITANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTAL	538
Qy	539	ALSCMADYGAPVPPFLAIGGKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLE	598
Db	539	ALSCMADYGAPVPPFLAIGGKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLE	598
Qy	599	ENRAFORWAGMFOVTFYAESLEDEINRTTAEDLPIFATSVIVFLYISLALGSYSWS	658
Db	599	ENRAFORWAGMFOVTFYAESLEDEINRTTAEDLPIFATSVIVFLYISLALGSYSWS	658
Qy	659	RVMVDSKATLGLGGVAVLVGAVMAAGPFSYLGIRSSVLIVQVPLVLSVGADNIFIV	718
Db	659	RVMVDSKATLGLGGVAVLVGAVMAAGPFSYLGIRSSVLIVQVPLVLSVGADNIFIV	718
Qy	719	LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEATCFELGALTMPAVRTFALTSGL	778
Db	719	LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEATCFELGALTMPAVRTFALTSGL	778
Qy	779	AVILDFLLQMSAFVALLSDSKQASRLDVCCKVQBELPPPGQEGELLGFFQKAYAP	838
Db	779	AVILDFLLQMSAFVALLSDSKQASRLDVCCKVQBELPPPGQEGELLGFFQKAYAP	838
Qy	839	FLHLWITRGVW----LLLFLALFGVSLYSNMCHISVGLDQELALPK	879
Db	839	FLHLWITRGVWVPSQLLFLALFGVSLYSNMCHISVGLDQELALPK	883
RESULT 18			
ID	AAW88445	AAW88445 standard; protein; 1278 AA.	
XX	XX		
AC	AAW88445;		
XX	XX		
DT	26-APR-1999	(first entry)	
XX	XX	Human NPC1 (Niemann-Pick type C disease) polypeptide.	
XX	XX		
KW	Niemann-Pick disease type C; NPC1 gene; human; diagnosis; therapy;		
KW	cholesterol; neurodegeneration.		
OS	Homo sapiens.		
PH	' Key	Location/Qualifiers	
FT	Peptide	1. .24	
FT	Protein	/note= "signal peptide"	
FT	Region	25. .1278	
FT	Region	/note= "signal peptide"	
FT	Region	55. .164	
FT	Region	/note= "conserved region"	
FT	Region	73. .94	
FT	Region	/note= "leucine heptad motif, or leucine zipper"	

FT	Peptide	1275. .1278	/note= "lysosome taretting motif"
XX	XX	WO9901555-A1.	
XX	XX	14-JAN-1999.	
XX	XX	02-JUL-1998;	
XX	XX	98WO-US013862.	
XX	XX	03-JUL-1997;	
XX	XX	97US-0051682P.	
XX	XX	(USSH) US DEPT HEALTH & HUMAN RESOURCES.	
XX	XX	Carstea ED, Tagle DA, Morris JA, Pentchev PG, Pavan WJ;	
XX	XX	Rosenfeld MA, Lofthus SK, Gu J;	
XX	XX	WPI; 1999-106056/09.	
XX	XX	N-PSDB; AAX06873.	
XX	XX	New isolated gene, NPC-1 - is associated with Niemann-Pick type C	
XX	XX	disease, used to develop products for the study, diagnosis and therapy of	
XX	XX	the disease.	
XX	XX	Claim 1; Page 56-60; 101pp; English.	
XX	XX	This polypeptide is encoded by the human NPC-1 gene (see AAX06873) that	
XX	XX	is associated with Niemann-Pick disease type C (NP-C). NPC1 when	
XX	XX	expressed in human cells derived from NP-C patients is able to correct	
XX	XX	the abnormal lysosomal cholesterol accumulation that is characteristic of	
XX	XX	such cells. Hydrophobicity studies and structural motif comparisons	
XX	XX	suggest the polypeptide is an integral membrane protein with as many as	
XX	XX	16 possible transmembrane regions. Database sequence comparisons reveal	
XX	XX	extensive identity/similarity to NP-C orthologues in mouse, yeast and	
XX	XX	Caenorhabditis elegans (see AAW88446-48). The provision of the human NPC1	
XX	XX	cDNA sequence enables methods of detecting the presence of mutations in	
XX	XX	the gene corresponding to the hNPC1 cDNA, and thereby facilitates the	
XX	XX	determination of whether an individual is an NP-C sufferer, or an NP-C	
XX	XX	carrier. The products can also be used for the study of NP-C and in	
XX	XX	therapeutic applications related to NP-C	
XX	XX	Sequence 1278 AA;	
XX	XX	Query Match 34.8%; Score 2402.5; DB 2; Length 1278;	
XX	XX	Best Local Similarity 39.8%; Pred. No. 1.8e-214;	
XX	XX	Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;	
Qy	7	RGLWLLALLLRLAQSEPTTTHQPCYCAFYDECG-----KNPELSGLMTLSNVCLS	59
Db	4	RGLALGLLLLLLCPAQVFSG-----SCVWYGEGLIAYGDKRYNCEYSG-----	46
Qy	60	NTPARKITGDHLLILQIKICPRLYTGTGNTQACCSAKQLVSLSEASLITKALLTRCPACSDN	119
Db	47	--PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDNLQLPLQLSRCPSCFYN	103
Qy	120	FVNHLCHNTCSNQSFLINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAESYDSCSRV	175
Db	104	LLNLFCELTCSPRQSQFLNVTATEDYDVPVTNQTKNVKELQYVYGQSPANAMYACRDV	163
Qy	176	RVPAATATLVGTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDIT-----FHLLEPGQA	228
Db	164	EAPSNDRKALGLLCKGRADA-CNATNWIYFMFNKNGQAPFTITPVFSDFPVH-----	215
Qy	229	VSGGIQPLNEGVARCNESQGGDDVATCSQDCAASCPAIARPO-----ALDSTFYLG	279
Db	216	---GMEPMNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWPTILGLDAMYIV	272
Qy	280	QMPGSLVLIILCSVFVAVVTILL-----VGFVAPARDKSKMVDPKKGTSL	326
Db	273	WITYMAFLLVFFGAPFAVWCKRYKRYVSEYTPIDSNIAFSV-NASDKG-----ASCC	324
Qy	327	DKLSFSTHTLLGQFFQGMGTWVASWPLTILVLSVTPVVALAAGLVFTELTTPDELWSAP	386
Db	325	DPVSAAFEGCLRLFRNGSPFCVRNPGCVIFFSLVPITACSSGLVFRVVTNPDVLSAP	384

Db	47	--PKPLPKGYDLVQELPCGFFPG-NVSLCCDVRQLQTLKDNQLQLPQLQFLSRCPSCFYN	103
Qy	120	FVNHLCHNTCSNOSLFINVTR----VAQLGAGQLPAVAYEAFYQHSFABQSYDSCSR	175
Db	104	LLNLFCELTCSQROFLNVTATEDYDVPNTQTKTNVKELOYYVGQSFANAMYNACRDV	163
Qy	176	RVPAAATLAVGTMCGVYGSALENAQRWLNFOGDTGNGLAFLADIT-----PHLEPQOA	228
Db	164	EAPSSNDKALGILLCKODADA-CNATNTIYEMFNKDNQGOAPFTITPVFSDFPVH-----	215
Qy	229	VSGIQPLNEGVARCNESGDDVATCSQDCDCAASCAPAJARPO-----ALDSTFYLG	279
Db	216	---GNEPMWNATKGCDESVDEWAPCSQDCSIVCGPKPQPPPPAPWTIILGLDAMYIM	272
Qy	280	OMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSKWDPKKGTSLS	326
Db	273	WITYMAFLLVFGAFAVCWKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC	324
Qy	327	DKLSFSTHTLLQFQGGTGWASWPLTILVLVSVVVALAAGLVFTLTTPDVELWAP	386
Db	325	DPVSAAFEGCLRLFRMGFCVRNPGCVIFFSLVFITACSSGLVFVRVTINPVDLWAP	384
Qy	387	NSQARSEKAFDQHGPPFPRTNQVILTAPNRSYRSDSLGPK-NFSGIILDLDLLELL	445
Db	385	SSQARLEKEYFDQHGPPFPRTQLIIRAPLTDKHIYQYPVGADVPFGPPLDIQLHQVL	444
Qy	446	ELQERLRLHVMSPEAQRNISQDICYAPLPDNTSLYDCCINSLLQYFQNNRTLLLLTA	505
Db	445	DLQIAIEN--ITASVDNETVTLQDICLAPLSYNT---NCTILSVLNYFQNSHVDHKK	499
Qy	506	NOTLMQOTSOVDKWHFLCANAPLTFKDGTALESMAADYCAPVFPFLAIGYKGKDY	565
Db	500	GDDFF---VYADYHTHFLVYCRAPASLNDTSLHDPCLGTFGPGVFPMLVGLGQDDQNY	556
Qy	566	EBAELMTFSLNYPAGDRLAQALWEAFLAEMRAFORMMAGMFQVFTTFAERSLEDEI	625
Db	557	NATALVITFPVNNYNDTELQRAQAWKEFINFKVKNY---PNLTISFTAERSIEDEL	613
Qy	626	NRTTAEPLPIFATYIVIFLYLISLALGSYSMSRSMVDSKATILGGVAVVIGAVMAWG	685
Db	614	NRESDSVFTWISYAIMELYISALGHKSKRLLVDSKVSIGIAGILIVLSSVACSIG	673
Qy	686	FSYLGIRSLVILQVPELVLSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAP	745
Db	674	VFSYIGLPLTLVIVIEPFLVLAAGVDNIFILVOAYORDERLQGETLQDLQGRVLGEVAP	733
Qy	746	SMLLCSLSEAI CFPLGALTPMPAVRTFALTSLGIAVLIDFELQMSAFVALLSLDSRQEAS	805
Db	734	SMELSFSEITVAFELGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDLKQEK	793
Qy	806	RLDVCCCKPQELPPPGQ-CEGILLGFPQKAYAPFLLHWITRGVVLILLFLALFGVLSYM	864
Db	794	RLDIFCVRGAEDGTSVQASESCLFRFFKNYSYSPILLKDMRPIVIAIFVGLSPFSIAVL	853
Qy	865	CHLSVGLDQELAPKOSLYLDYFLNRYPEVCAPVYFVTTILGYNSSBAGMAICSSAG	924
Db	854	NKVDIGLDQSLPDDSDMYDPKSIISQYLHAGPPVTVLEBHDITSSKGQNNVCGMG	913
Qy	925	CNNPSTFKTQYATEFPEQSYLAIPASSWVDDFIDMLTP--SSCCRLYISGPNKDFCPST	983
Db	914	CNNDLSVQOIFNAQLDNYTRIGFAPSWDIDDYFDMVKPQSSCCRV---DNITDQCNAS	970
Qy	984	VNSLNLCKNCSITM-GSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAYSTSVN--LTS	1040
Db	971	VDPACVR-CRPLTPGKORPQGGDFMRFLFELMSDNPNPKCGKGGHAAYSSAVNILLGH	1029
Qy	1041	DGOVLASREWAYHKPLKNSDYTEALRAARELANITADLRKVPGTDPAPVEVFPYITNV	1100
Db	1030	GTRVGATYFMYHTVLTQTSADFDALKAKALIASNT-ETWINGS--ARVFPYSYFVY	1086
Qy	1101	FYEQLATLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLISIVMLVDTVGVFMALWD	1160
Db	1087	FYEQLVLIIDTIFNLGVSIGALFLVTWVLLGCELWSAVIMCATIAMVLVNMFGVWMLMG	1146
Qy	1161	ISYNAVSLINLVSAGMSVEFVSHITRSPAITSTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1147	ISUNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK	1206
Qy	1221	LPGILVLGLAKAQLIQIFFPRLNLLTLLGLLHGLVFLPVLVSVYGVDPVNPA	1272
Db	1207	FGIVVLAFAKSQIFQIFPYRMYLAMVLLGATHGLIFLPVLLSYIGPSVNKA	1258
RESULT 20			
ADQ39879			
ID	ADQ39879 standard; protein; 1278 AA.		
XX	ADQ39879;		
AC	ADQ39879;		
XX	18-NOV-2004 (first entry)		
DT	Human myocardial infarction-associated gene derived protein, SEQ ID 1542.		
XX	Myocardial infarction; detection; single nucleotide polymorphism; SNP;		
DE	cardiant; gene therapy; human.		
XX	Homo sapiens.		
OS	WO2004058052-A2.		
XX	15-JUL-2004.		
PD	22-DEC-2003; 2003WO-US040978.		
XX	20-DEC-2002; 2002US-0434778P.		
PR	10-MAR-2003; 2003US-0453135P.		
PR	30-APR-2003; 2003US-0466412P.		
PR	23-SEP-2003; 2003US-0504955P.		
XX	(APPL-) APPLERA CORP.		
XX	Cargill M, Devlin J, Iakoubova O;		
PI	WPI; 2004-533949/51.		
XX	N-PSDB; ADQ39051.		
DR	Identifying an individual who has an altered risk for developing myocardial infarction by detecting a single nucleotide polymorphism in the individual's nucleic acids.		
XX	Claim 10; SEQ ID NO 1542; 145pp; English.		
PS	The invention relates to a novel method for identifying an individual who has an altered risk for developing myocardial infarction. The method comprises detecting a single nucleotide polymorphism (SNP) in any one of the nucleotide sequences given in the specification in the individual's nucleic acids, where the presence of the SNP is correlated with an altered risk for myocardial infarction in the individual. The invention further comprises: an isolated nucleic acid molecule comprising at least 8 contiguous nucleotides where one of the nucleotides is an SNP given in the specification or its complement and encoding any one of the amino acid sequences given in the specification; an isolated polypeptide comprising an amino acid sequence given in the specification; an antibody that specifically binds to the polypeptide or its antigen-binding fragment; an amplified polynucleotide containing an SNP given in the specification and which is between about 16 and 1000 nucleotides in length; a kit for detecting an SNP in a nucleic acid, comprising the polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a nucleic acid molecule; a method of detecting a variant polypeptide; and a method for identifying an agent useful in treating or preventing myocardial infarction. The novel detection method has cardiant activity. The nucleic acids of the invention may be used in gene therapy. The method is useful in identifying an individual who has an increased or decreased risk for developing myocardial infarction and for preparing a composition for treating or preventing myocardial infarction. This composition represents the protein of a human myocardial infarction-		

CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX * Sequence 1278 AA;

Query Match 34.8%; Score 2402.5; DB 8; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.8e-214;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLAQSPPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVSQ-----SCVYGEGLIAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLILLOKICPRILYTGNTQACCSAKQVLSLEASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDLQLPLQFLSRCPSCFYN 103
QY 120 FVNLCHTNCTSPNQSFLINWTR-----VAQLGAGQLPAVVAYEAFYQHSFABQSYDSCSRV 175
DB 104 LNLFLCELTCSPRQSQFLNVTATEDYVDFVTNQTKTNKVELQYVVGQSPANAMYNACRDV 163
QY 176 RYPAATATLAVGTMCGVYGSAICNAQRWLNFGQDGTGNGLAPLDIT-----PHLEPGQA 228
DB 164 EAPSNDRKALGLGCKGADA-CNATNWIEMFNKONGQAPFTITPVFSDFPVH----- 215
QY 229 VSGSIGQLMEGVARNCSQDDVATCSQDCQCAASCPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGDSEVDEVTAPCSQDCSIVCGPKPQPPPPAPWTLILGDAMYIM 272
QY 280 QMPGSLVLIILCSVPVAVTILL-----VGRVPAPARDKSMVDPKKGTSLS 326
DB 273 WITYMAFLLVFGAFVAVCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSSTHTLLQFGQGWGTWASPLTILVLSVPVALAAGLVFTELTTDPVELWSAP 386
DB 325 DVSAAFEGLRLRFTWGSFCVRNPGCVIFPSLVFITACSGLVFVRVTTNPFVLSAP 384
QY 387 NSQARSEKAFHQHGFPPRTNQVILTAPNRGSRYDLSLLGPK-NFSGILDLDLLELL 445
DB 385 SSOARLEKEYFOHGFPPFTQLIIRAPLTKHIYQPYPSGADVPFGPPLDIQILHQL 444
QY 446 ELQERLRLHQLVMSPEARQNISLQDICYAPLNPNTSLYDCCINSLLQYPPQNNRTLLTPA 505
DB 445 DLQIAIEN--ITASVDNETVTQLDCLAPLSPYNT---NCTILSVLNYFQNGSHVLDHKK 499
QY 506 NQTLMGQTSQVQDKHFLYCANAPLTFKDGTLALSCMADYCAPVFPFLAIGYKGDYS 565
DB 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLHDPCLGTGCGVPPVPLVLGGYDDQNYN 556
QY 566 EBAALIMTFLSNYPAGDPRLAQAKLWBEAFLEEMRAFORMAGMFQVTTFAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKQRAQWEKEFINFVNKYKN---PNLTISFAERSLEDEI 613
QY 626 NRTTAEDELFPATSYIVIFLYLSLALGYSYSSRVMSVDSKATILGGLGVAVVLGAYMAAG 685
DB 614 NRESDSDVFTVISYAINFLYLSLALGHKSCRELLVDSKSLGIAGIILVLSVACSIG 673
QY 686 PFYSYLGIRSLVILQVPPVLVSVGADNIFIFVLEYQRLPRRPGPREVHIQALGRVAP 745
DB 674 VFSYGLPLTLVIEVPEPLVAVGVDNIFILVQAYQRDERLQGETLDQQLGRVLGEVAP 733
QY 746 SMLCGLSGLAICFFLGALTPMAVTEALTSLGLAVLDFLQMSAPVALLSDSKROAS 805
DB 734 SMLFSFSFTVAFFGLGSLVMAVHTFSLFAGLAVFIDPLQLITCFVSLGLGDIKQKQKN 793
QY 806 RLDDVCCVKPQBLPPPGQ-GEGLLGGFFQKAYAPFLHWTIRGWVLLALFLALFGVSLYSM 864
DB 794 RLDDICVVRGAEDGTSVQASECLPRFFKNYSPLLLKDMWMPVIAIFVGLVLSIAVL 853
QY 865 CHISVGLDQELAPKDSYLLDYLFLNRYFVVGAPVYFVTTLGYNFPSSBAGNAICSSAG 924
DB 854 NKVDIGLQSLMPDSDVMVDYFKSITQYLHAGPPVYFVLEERHDTYTSKQGMVCGMG 913

QY 925 CNFSTFTQKIQVATBPPEQSYLAIPASSWVDWDFIDWLTTP-SSCCRLYISGPNKDKFCPST 983
DB 914 CNNDLSLVQIIFNAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNCNLCNMSIT-MGSVVRPSVEQHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VYDPACVR-CRPLTPEGKORPQGDPMRFLPMLFSDNPNPKCGKGAAYSSAVNILLGH 1029
QY 1041 DQCVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYTTINV 1100
DB 1030 GTRVGATYPTWTHVLOTSSADFDALKKARLIASNT-ETMGINGS--AYRFPPIYVFPV 1086
QY 1101 FYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSIVMILVDTVGFNALMD 1160
DB 1087 FYEQYLTITIDDTIFNLGVSLGAIFLVTWVLLGCELWSAIVMCATIAMVLVNMFGVWLWG 1146
QY 1161 ISYNAVSLINLVSAGVMSVEFVSHITRSFAITKPTWLERAKETISMGSAVPAGVAMTN 1220
DB 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRVERABEALAHMGSSVPSGITLTK 1206
QY 1221 LPGILVLGLAKAQLIQQPFRNLNLLITLGLLHGLVFLPVLISYVCGPDVNP 1272
DB 1207 FGGIVVLAPAKSQIFQIFFRMLAWVLGATHGLIFLPLVLLSYIGPSVNKA 1258

RESULT 21

ADU06723

ID ADU06723 standard; protein; 1278 AA.

XX AC ADU06723;

XX DT 27-JAN-2005 (first entry)

XX DE Novel bronchial cancer-associated human protein SeqID949.

XX KW bronchial cancer; cytostatic; tumour-associated protein;

XX KW cancer detection; metastasis; tumour; human.

XX OS Homo sapiens.

XX PN DE10316701-A1.

XX PD 04-NOV-2004.

XX PF 09-APR-2003; 2003DB-01016701.

XX PR 09-APR-2003; 2003DB-01016701.

XX PA (HINZ/) HINZMANN B.

XX PA (HERM/) HERMANN K.

XX PA (CAST/) HEIDEN CASTANOS-VELEZ E.

XX PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;

XX PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pillarsky C;

XX DR WPI: 2004-786403/78.

XX DR N-PSDB; ADU06236.

XX PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial cancer and in screening for therapeutic and diagnostic agents.

XX PS Claim 2; SEQ ID NO 949; 1381pp; German.

XX CC This invention relates to a novel isolated nucleic acid associated with bronchial cancer comprising 489 defined sequences given in the specification. The invention may be useful for the production of compounds with a cytostatic activity through the inhibition of expression or activity of tumour-associated proteins. The novel DNA sequences and the proteins/peptides encoded by them are used for detecting bronchial cancer or determining the risk of developing it and to screen for specific binding partners of the DNA or protein sequences, where the binding partners are potentially useful as agents for treating or

CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a protein
CC encoded by a novel bronchial cancer-associated human gene sequence of the
CC invention.

XX SQ Sequence 1278 AA;

Query Match 34.8%; Score 2402.5; DB 8; Length 1278;
Best Local Similarity 39.8%; Pred No. 1.8e-214;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
QY 7 RGMWALLRLRLAQSEPTTHIQGYCAPYBECG-----KNPELSGLMTLSNVSLIS 59
DB 4 RGLALGULLLLCPAQFSQ-----SCVMYGCBIAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHILLQKICPRLYTGPTQACCSAKOLVSLASISITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFFG--NVSLCCDVRQLQTLKDNLQLPLQFUSRCPCFYN 103
QY 120 FYNLCHNTCSNQSFLINVT---VAOLGAGQLPAVVAYEAFYQHSFAEQSYDCSRV 175
DB 104 LNLFLCELTCSPRQSGFLNVTATEDYDVPVNTQTNVKELYYVGQSFANMYNACRDV 163
QY 176 RVPAAATLAVGTGCMGYGSGALCNAQRWLNFOGDTGNGLAPLDIT-----PHLEPQQA 228
DB 164 EAPSSNDKALGLLCKGADA-CNATNWIEMFNKNGQAPPTITPVSDFPVH----- 215
QY 229 VSGIOPLEGVARNCHESOGDDVATCSODCAASCPAIAPO-----ALDSTFYLG 279
DB 216 --GMEPMNATKGCDSDEVTAPCCQDCSIVCGKFPQPPPPPPAPWTILGLDAMYIM 272
QY 280 QMPSGLVLIILCSFPAVVTILL-----VGRVAPARDKSMVDPKKGTSLIS 326
DB 273 WITYMAFLVFFCAFAVWCYKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSTFHTLLGQFFGQGTWASWPLTILVLSVIVVVALAAGLVPTELTTPDVLWSAP 386
DB 325 DPVSAAPFEGCLRLFRWGSFCVRNPGCVIFFSLVFITACSGSLVFRVVTTPVDLWSAP 384
QY 387 NSQARSEKAFHQHGFEPFTNQVILTAPNRSRYRVSLLGPK-NFSGILLDLILLELL 445
DB 385 SSOARLEKEFYDQHGFEPFTEQLIIRAPLUTDKHIYQYPFGADVPFGPLDIQILHQVL 444
QY 446 ELQERLRLQVMSPEAQRLISLOICVAPLNPONTSLYDCCINSLLQYFQNNRTLLLLPA 505
DB 445 DLQIAIEN--ITASYNETVTLQDLCIAPLSPYNT--NCTILSVLNYFQNSHSLVDHKK 499
QY 506 NQTLMGQTSQVNDKDHPLCANAPLTFKGTOTALALSCMADYGAVPFPFLAIGYKGDYS 565
DB 500 GDDPFF--VYADYHTFLYCVRAPASLNDTSLLDPLCLGTGFGVPFPLVLGGYDDQNYN 556
QY 566 EAEALIMTFSLNPNYPAGDPRLAOKLWEEAFLEEMRAFORRMAGMVFVTFATERSLEDEI 625
DB 557 NATALVTFFPNVNYNDTEKLQRAQWKEKFINFVKNKYL---PNLTISTAERSIEDEL 613
QY 626 NRTTAEDLPITFATSYIYIFLYISALGSYSWSRWVMSKATLGLGAVAVVLGAVMAAMG 685
DB 614 NRESDSVFTVVISYAIMFLYISALGHKISCRLLVDSKVSLGIAGILIVSSVACSLG 673
QY 686 FFSYLGTRSSLIVLQVVPFVLVSGADNIIIFVLEYQRLPRRGPPEVHIGRALGRVAP 745
DB 674 VFSYIGLPLTILVIEVIFPLVAVGVNDNIFILVQAYQORDERLQGETLDQQLGRVLGEVAP 733
QY 746 SMLLCSLSEATCFPLGALTWPAVRTFALTSGTAVILDLOMSAFVALLSLDSKQEQAS 805
DB 734 SMFLSSFSSETVAFPLGLALSNVPAVHTSFLAGLAVDFIDLLQITCVFSLGLGLDKRQKN 793
QY 806 RLDVCCCKVKPQLPFPQ--GEGLLLGFFQKAYAPFLHMTIRGVVLLLFALFGVSLYSM 864
DB 794 RLDIFCCVGAEDGTSVQASBSCLFRFFKNSYSPLLKDWMRPTVIAIFVGLVSFSAVL 853

QY 865 CHLSVGLDQELALPKDSYLLDYFLNRYPEVGCAPVYVFTTLTGLNFSSEAGMNAICSSAG 924
DB 854 NKVDIGLDQSLSPDDSYWVDYFKSISQYLAHAGPVPYFVLEEGHDYTSKQGNVCCGMG 913
QY 925 CNNFSTQKIQYATEPEQSYLAIPASSWVDDFIDMLTP--SSCRLYISGPNKDKPCPST 983
DB 914 CNNDLSVQQIFNAQAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNCSIT-MGSVPSVSEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VVDPACVYR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGGHAAYSASVAILLGH 1029
QY 1041 DGOVLASRFRMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFEVFPYTTIN 1100
DB 1030 GTRVGATYFTYHTVLTQTSADFIDALKARLIASNT-ETMGINGS--AYRVFPYSVFYV 1086
QY 1101 FYEQYLTLPPEGLPMLSLCLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTVDFGMALWD 1160
DB 1087 FYEQYLTIIODTIFNLGVSGLAIFLVTMVLGCELWSAVIMCATIAMVLNMFVGMWLMG 1146
QY 1161 ISYNAVSLINLVSAGVMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVPAGVAMTN 1220
DB 1147 ISLNAVSLVNLVMSCGISVBFCSHITRAFTVSMKGSVRABEALAHMGSSVFSGIILTK 1206
QY 1221 LPGLILVLGLAKAQIQLIFFFRNLILLITLLGLLHGLVFLPVLISYVGPDVNPA 1272
DB 1207 FGGIVLWLAFAKSOIQFYFMYLAWLLGATHGLIFLPLVLSYIGPSVKA 1258
RESULT 22
AAW88446
ID AAW88446 standard; protein; 1319 AA.
XX AC AAW88446;
XX DT 26-APR-1999 (first entry)
XX DE Mouse NPC1 orthologue.
XX XX Niemann-Pick disease type C; NPC1 gene; mouse; orthologue; diagnosis;
KW therapy; cholesterol; neurodegeneration.
XX OS Mus sp.
XX PN WO9901555-A1.
XX PD 14-JAN-1999.
XX PF 02-JUL-1998; 98WO-US013862.
XX PR 03-JUL-1997; 97US-0051682P.
XX PA (USSH) US DEPT HEALTH & HUMAN RESOURCES.
XX PI Carstee ED, Tagle DA, Morris JA, Pentchev PG, Pavan WJ;
PI Rosenfeld MA, Loftus SK, Gu J;
XX DR WPI; 1999-106056/09.
XX DR N-PSDB; AAX06874.
PT New isolated gene, NPC-1 - is associated with Niemann-Pick type C
PT disease, used to develop products for the study, diagnosis and therapy of
PT the disease.
XX Claim 1; Page 66-70; 101pp; English.
XX This polypeptide comprises the murine orthologue of human NPC-1 (see
CC AAW99445) that is associated with Niemann-Pick disease type C (NP-C). The
CC amino acid sequence, predicted from an isolated cDNA clone (see
CC AAX06874), comprises a putative N-terminal signal peptide followed by a
CC domain that is unique to the NPC1 orthologues (in mouse, human,
CC Caenorhabditis elegans and Saccharomyces cerevisiae, see AAW88445-48),

CC resources for understanding the role of NPC1 in intracellular cholesterol
CC homeostasis and in the aetiology of neurodegeneration in NP-C disease.
CC The provision of the human NPC1 cDNA sequence (see AAX06873) enables
CC methods of detecting the presence of mutations in the hNPC1 gene, and
CC thereby facilitates the determination of whether an individual is an NP-C
CC sufferer or carrier
XX

SQ Sequence 1296 AA;
Query Match 15.6%; Score 1078; DB 2; Length 1296;
Best Local Similarity 24.0%; Pred. No. 2.2e-90;
Matches 321; Conservative 237; Mismatches 492; Indels 288; Gaps 41;

QY 76 KICPRLTYGPNTOACSAKQLVSLRASISITKALLTRCPACSDNFVNLHCHTCSPNQSL 135
DB 63 EPCPHLLTGDN-KLCTPSQABGLTKQAQAARHILGRCPSCFDNFALWCEFTCSNPOOD 121
QY 136 FINTVFAQL--GAGOLPAVAYEAF-----YQHS--PAEQSYDSCSRVRPAAATLAVG 186
DB 122 FVSISEMKEPIEKKEGTPYQAEAYVNTVEYRLSTDFAEGMFSSCKDVTFGGQPALAV- 180
QY 187 TWCVVYGSALCAQNALNFGQDTGNGL-APLDITFHLLEPGQAVSGIQP-LNEGVARCN 244
DB 181 -MC---TSPCTLTNWLFIQTQNLNIPHTKLLYDPIKTPPSDRSTYNNVNFVFTGCD 236
QY 245 ESGDDVATCSCQDCAASCPAIPALDSTFYLGOM-----PGSLVLIILCSVFAVVT 299
DB 237 KSAVWGPACTSEC--NKEEYANLIDLDDGKTSQGTQCNVHGACINLFIWMLAFITGSLAV 294
QY 300 ILLVGF-----RVAPARDKSKMVDPKKGTSLSDKLSFSFTLLLGQFF 341
DB 295 LLCVGFTSYSDYTNLRQTSGEESPRNRIK-----RTGAWI-----HNFMENNA 342
QY 342 QCGMTVWASWPLTILVLSVIVVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHF 401
DB 343 RDIGMWAGRNPKSHFFIGCAVLIFCLPGMIYHKESTNVVDMWSSPRRARQBEWVFANF 402
QY 402 GPFRFNQVLLTAPNRSSRYDLSLLGPKNFSGILDLDLLLELLEQLERLHLQWSPEA 461
DB 403 GRPQRYQQIMLL--SHRDFQSSGKLYCP-----VFHKDIFEEFLFDILNAIKNISTQDSG 455
QY 462 QRNISLQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLTANOT-----L 509
DB 456 -RIITLDDVCYRPMGPG-----YCLIMSPNYFQGNKEHLDKMSKEETVSEDDDAFYF 510
QY 510 MGQTSQVDKWDHFLYCANAPLTPKQGTALALSCWADYGPVFPFLAIGGYKGDYSEAA 569
DB 511 SSEATTDEWNNHMAACIDQPMQSK--TKSGLSCWGTGYPGSPAPNM-VFGKNSTNHQAANS 567
QY 570 LIMTFSLNYPAGDPRLAQAQKLWEAEFLWEAFQRMAGMFOVFTFAERSLEDEINRTT 629
DB 568 IMMTILVTO--RTEPEIQKAELEKEKPLKFKCYREKSPKVI-FSFMARSITDEIENDA 624
QY 630 AEDLPIFATSVIVIFLYISALASY-----SSWSRVWVDSKATILGLGVAVLGAVMAAM 684
DB 625 KDEIVTVVIALAFILGIVTFSLGRYFVCENQLWS-ILVHSR----- 664
QY 685 GFPSYLGIRSLVLQVVPFLVLSVGADNIFIVLEYQRLPRRPGEPREVIHIGALGRVA 744
DB 665 GF----- 666
QY 745 PSMLLCSLSEACFFLGNALTPMPAVRTFALTSLGLAVILDLLQWSAPVALLSDLSKROEA 804
DB 667 -----TDLPAIRTFCLYAGLAVLIDVVLHCTIFLALFVWDYQRELN 707
QY 805 SRLDVCCCKVQELPPGQGEGL-----LLG-----FFQKAYAPFLHLWITRGVV 849
DB 708 G-----KP-EFFFPFQIKDLGLAGVILIGQRATDTFTMQFFHFQVAPFLWHRMTRIIT 758
QY 850 LLLFLALFGVSLYSNMCHSVGLDELALPKDSYLLDYFLFLNRYFVEGAPVYFVTTLGYN 909
DB 759 GIIFIASPITTVILSSKISVSGFDQSMAPTEKSYISTHFRYLDKDFDVGPPVFFVTDGELD 818

QY 910 FSSEAGMNAICSSAGCWNFSFTQKIQVATEPPEBQSYLAIPASSWVDVDFIDWLT-PSSCCR 968
DB 819 WHPDPQNKFCCTPGCSDTSTFGNIMYAVGHTETQYLSGEMWNIDWYLEWISKSPCK 878
QY 969 LYISGPNKDFCPSSTVNSL-----NCLKCMISITWGSVRPSVEQFHKYLPWFLNDRPNIC 1024
DB 879 VYVHDPN--TFCSTNRNKSALDDKACRTCMDF----- 908
QY 1025 PKGLAAYSTSVNLTDGQVLASRFMAYHKLPL--KNSQDYTEALRAARELAANITADLRK 1082
DB 909 --DGRASFKAISTSTSGRIQASQFMTHFKKLSISNSDFDKAMDYARWYSRRLERSI-- 964
QY 1083 VPGTDPAFEPYPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNL 1142
DB 965 ----DDTAHFVAFYSKIFPFYEQYSTIMPILTTQLFITVVGFGIICVTGLGIDVKAACAV 1020
QY 1143 LSIWMLIVDTVGFALWDISYNAVSLNLINLNAVCMGVSEFVSHITRSEFAISTKPTWLERAK 1202
DB 1021 ICQVS-----NYFHIWSS--GILIEFVNVVLKGYACSLURQAKORAE 1060
QY 1203 EATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLITLLGLHLGLVFLPVIL 1262
DB 1061 STVGSIGPILSGPVVTWAGSTMFLSGAHLQIITVYFFKFLITIVSSAVHALIILPILL 1120
QY 1263 SYVGPDVNPALALEQKRAEAAVAAVMVASCNHPSPRVSTADNIYV-----HSEGS 1314
DB 1121 AFGSRGHGSSETSTNDNDQHDACVLS--PTAESHSINVEEGILNRPSLDDASHILDPL 1178
QY 1315 IKGAGALSNFLPNNRQF 1332
DB 1179 LKAEGGIDKAI---GRDF 1193

RESULT 28

AAM80153
ID AAM80153 standard; protein; 229 AA.

AC AAM80153;

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 3799.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

PN WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK53286.


```

PF 22-MAR-2001; 2001WO-JP002279.
XX
XX
PR 24-MAR-2000; 2000JP-00088595.
XX
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
XX
PI Taniyama Y, Kita S, Komiyama T;
XX
XX WPI; 2001-611501/70.
XX
XX N-PSDB; AAI66914, AAI66931.
XX
XX
XX New sterol-sensing domain-containing protein for diagnosing and screening
XX candidate compounds in drug development for diabetes, obesity, cancer,
XX arteriosclerosis, hyperlipidemia and neurodegenerative disorders.
XX
XX Claim 6; Page 113-114; 171pp; Japanese.
XX
XX The invention provides a novel SSD (sterol-sensing domain)-containing
XX protein. The protein originates from human liver, human testis or human
XX brain. The protein can be expressed by standard recombinant methodology.
XX The proteins, encoded DNAs and antibodies are useful in diagnosis and
XX screening candidate compounds in drug development for diabetes, obesity,
XX cancer, arteriosclerosis, hyperlipemia, neurodegenerative disorders such
XX as Alzheimer's disease and neural disorders. The present sequence
XX represents a human SSP1 SSD domain
XX
XX Sequence 194 AA;
SQ
Query Match 13.6%; Score 942; DB 4; Length 194;
Best Local Similarity 99.5%; Pred. No. 5e-79;
Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 615 FTAERSLEINRTAEDLPFATSVIVIFLYISALGSYSWSRMVDSKATLGLGVA 674
DB 1 FWAERSLEINRTAEDLPFATSVIVIFLYISALGSYSWSRMVDSKATLGLGVA 60
QY 675 VVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREV 734
DB 61 VVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREV 120
QY 735 HIGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVAL 794
DB 121 HIGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVAL 180
QY 795 LSLDSKRQASRLD 808
DB 181 LSLDSKRQASRLD 194
RESULT 31
AAR75375
ID AAR75375 standard; protein; 1447 AA.
XX
XX AAR75375;
AC
XX
XX 30-JUL-1996 (first entry)
DT
XX
XX Human patched protein.
DE
XX
XX Patched gene; PTC protein; embryo development; cellular regulation;
KW signal transduction; ligand; antibody; hedgehog protein.
XX
XX Homo sapiens.
OS
XX
XX WO9611260-A1.
PN
XX
XX 18-APR-1996.
PD
XX
XX 06-OCT-1995; 95WO-US013233.
PF
XX
XX 07-OCT-1994; 94US-00319745.
PR
XX
XX (STRD ) UNIV LELAND STANFORD JUNIOR.
FA

```

```

XX Scott MP, Goodrich LV, Johnson RL;
PI
XX
XX WPI; 1996-209842/21.
XX
XX N-PSDB; AAT14220.
XX
XX DNA encoding patched protein other than Drosophila patched protein - used
XX to produce antibodies which detect or inhibit patched protein ligand
XX signal transduction in cells.
XX
XX Example; Page 56-61; 70pp; English.
XX
XX Human patched protein (PTC) (AAR75375) was identified as the product of a
XX cDNA clone (AAT14220) derived from human lung. It has 96% identity and
XX 98% similarity to mouse PTC. PTC protein has been proposed as a receptor
XX for hedgehog protein on the basis of genetic experiments in flies. Human
XX PTC protein can be obtained in large amounts by expression of the cDNA clone
XX in transformed host cells. It can be used to screen for agonists and
XX antagonists to isolate its ligand, partic. Sonic hedgehog, to assay for
XX the transcription of ptc mRNA and to raise antibodies
XX
XX Sequence 1447 AA;
SQ
Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 2.4e-50;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY 225 PGQAVSGIQIPLNEGVARCNESQGDVVATCSQDCAACPAIARPQALDSTFVLGMPGS 284
DB 24 PGRPAGGRRRTTGLRAAAPDRD-----YLRPSYCDAAFALEQI--- 65
QY 285 LVLIILILCSVPFVVTILLVGRVPAPARDKSMVDPK-----KGTSLSKLSFSTHTLL 337
DB 66 -----SKGATGRKAPLWRAKFORLLFKGCVIQKNC 98
QY 338 GQFFQGWGTWASWPLTILVSLVIPVVALAAGLVFTTELTPDVLWAPNSQARSEKAFH 397
DB 99 GKF-----LVVGLLIFGAFVGLKAANLETNVEELWVEVGGRVSRRLNYT 143
QY 398 DQHFGPFRTY-QVILTAPNRSSVRYDSLLGPKNFGSILDLDLLELLE--LQERLRLH 454
DB 144 RQKIGEEAMFNPQLMIQTPKEEG-----ANVLTEALLQHLSALQASRVHV 190
QY 455 QVWSPEAQRNLSIQDICVAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 191 YMYN-----RQWKLEHLCKYSGELITETGYMDQIIEYLVPCLIITPLDCFWEGAKLQSGTA 246
QY 506 NQTLMGQ-----TSQVDWKDHFY-----CAN----- 527
DB 247 --YLLGKPLRWNTNFDPLFLEELKINYOQVDSWBEMLNKAEGVGHYMDRPLCLNPADPDC 304
QY 528 -----APL-----TFKDG-----ALALSCWA 544
DB 305 PATAPNKNSTKPLDMALVNGGCHLSRKYHWQBELIVGGTVKNSTGKLVSAHALQTMF 364
QY 545 DYGAIVFPFLAIGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLEMRAFO 604
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE---DKAAAILAEAMQRTVVEVHOSV 411
QY 605 RMAGMFOVTFPAERSLEDEINRTAEDLPFATSVIVIFLYISALGSYSWSRMVDS 664
DB 412 AQNSTQKVLSTFT--TLDDILKSPSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
QY 665 KATLGLGVAVVLGAVMAAMGFYSYLGIRSSVILQVVPFLVLSVGADNIFIFVLEYQRL 724
DB 466 QGAVGLAGVLLVALSVAGLGLCSLIGISFNAAATTVLPFLALGVGVDDVFLAHAFSET 525
QY 725 PRPGEPREVHTGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTSLGLAVILO 784
DB 526 GQNKRIFFEDRTGECLKRTGASVALTSISNVTAFFMAALIPFALRAFSLQAAVVVVFNF 585
QY 785 LLQMSAFVALLSLDSKRQASRLDVCCC-----VKQDEL-----PPPG 822

```

Db 586 AMVLLIPAILSMYDLYRREDRLDIFCCFTSPCVSRVIOVEPQAYTDTHNTRYSPPPY 645
 QY 823 QCEGL----- 827
 Db 645 SSHSPAHEQITMQSTVQLRTEVDPHVHYVYTAEPRESEISVQPVTVTQDTLSCQSPST 705
 QY 828 -----LLGFPKAYAPFILLHMTIGVGLLLFLALFGVSL 861
 Db 706 SSTRDILSQFSDSLHCLBPCTKWTLSGFAEKHAPFLKPKAKVWVFLFLGLLGVSL 765
 QY 862 YSMCHISVGLDQELAPKDSYLLDYFLNRRFEVCAVYFVTLGYNFSSSNGMAICS 921
 Db 766 YGTRVRDGLDLTDIVRETRYEDFAAQKFYF-----YNN----- 803
 QY 922 SAGCNNSFTQKIYAT-----EPQESYLAIPASS-----WVDDFDIML----- 961
 Db 804 -----YIVTQADYNIQHLHYDLHRSFNKVMLEENKQLPKWMLHYFRDMLQGLQD 857
 QY 962 -----TPSS-----CCRYIYSGPNKDFCPTVNSLNCNKMSITMGS 1000
 Db 858 AFDSDWETGKIMPNNYKNGSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
 QY 1001 VRPSVEQPHYL-PWFLND-----RPNKCPKGLAAY--STSVNLTSDQ 1043
 Db 914 INPSA--FYIYLTAMVNSNDPVAYASQANIRPHRPEWHDK--ADYMPETRLRIPAAEP 968
 QY 1044 VLASRFMAHYKPKNSQDYTEALRAARELAANITA-DLRKVPCTDPAPFVFPVTTITNVFY 1102
 Db 969 IEYAQPPFYGLNGRLDTSDFEALKEKVTICSNYTSGLSSYPNG-----YFF-----LFW 1018
 QY 1103 RQYTLPLPGLMWSLCVPTFAVSCLLGLDLRSLGLMLLSIVMLVDTVGFMALWDIS 1162
 Db 1019 EQYIGLRHMLLLFISVVLACTFLVCVFLNPNWTAGII-VVMLALMTVELFGWMLGIGK 1077
 QY 1163 YNAVSLINLVSAVGSVEFVSHITRSP--AISTKPTWLERAKENTISMGSAVPAVMTN 1220
 Db 1078 LSAVPVILLIISVGIGVEFVTHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
 QY 1221 LPGILVLGLAKAQIQIPFRMLNLITLGLLHGLVFLPVILSYVG--PDVNPALALEQ- 1277
 Db 1134 LGLVLMAGSEDFIVRYFVAVLAILTILGVNLGLVLLPVLLSPFGPFGPEVSPANGLRL 1193
 QY 1278 --KRAEEAAMVAVMASCNPHSRVSTADNIYVNHSPGSIK 1317
 Db 1194 PTPSPPPSVVRFAAMPCHGTH--SGSDSDSEYSSQTTVSG 1233

RESULT 32
 AAW52200
 ID AAW52200 standard; protein; 1447 AA.
 XX
 AC AAW52200;
 XX
 DT 25-JUN-1998 (first entry)
 XX
 DE Human patched (ptc) protein.
 XX
 KW Patched protein; ptc; cancer; tumour suppressor; cell adhesion promoter;
 KW wound healing; ageing; human.
 XX
 OS Homo sapiens.
 XX
 PN WO9745541-A2.
 XX
 PD 04-DEC-1997.
 XX
 PF 02-JUN-1997; 97MO-US009553.
 XX
 PR 31-MAY-1996; 96US-00656055.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PA (REGC) UNIV CALIFORNIA.
 XX

PI Scott MP, Goodrich LV, Johnson RL, Epstein E, Oro A;
 XX WPI: 1998-032648/03.
 DR N-PSDB; AAV21590.

PT Patched protein other than Drosophila melanogaster patched protein - used
 PT for characterising the phenotype of a tumour.

XX Claim 4; Page 76-80; 86pp; English.

CC This is a human patched (ptc) protein. The encoding DNA can be used to
 CC construct an expression cassette comprising an altered patch or hedgehog
 CC gene. The expression cassette comprises a nucleic acid encoding a patched
 CC protein other than a Drosophila melanogaster patched protein, or fragment
 CC of at least 12 nucleotides in length, as other than an intact chromosome
 CC under transcriptional control of a transcriptional initiation region, and
 CC a transcriptional termination region, both functional in an expression
 CC host. A genetically engineered mammalian cell comprising this expression
 CC cassette as an extrachromosomal element or integrated into the genome of
 CC the cell can be predisposed to develop basal cell carcinoma as a result
 CC of the transfection. By analysing DNA, functional analysis of patched
 CC protein function, or by detecting antibody binding to abnormal patched
 CC protein, a genetic predisposition to developmental abnormalities and
 CC cancer can be diagnosed. This analysis can also be used for
 CC characterising the phenotype of a tumour, particularly a carcinoma,
 CC especially a basal cell carcinoma. The methods can also be used for
 CC characterising transitional cell carcinoma of the bladder, meningiomas
 CC medulloblastomas, etc. The modified cells comprising the expression
 CC cassette can be used to determine the role of different exons of the
 CC patched gene in oncogenesis, signal transduction, etc. Transgenic animal
 CC models created from these cells can be used as animal models for
 CC carcinomas of the skin. The patched protein of mosquito, butterfly or
 CC beetle or alternatively, a mammalian patched protein of human or mouse
 CC can be used to identify ligands or substrates that bind to, modulate, or
 CC mimic the action of patched gene. These agents could be used as tumour
 CC suppressors, cell adhesion promoters (e.g. in wound healing and ageing)
 XX Sequence 1447 AA;

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
 Best Local Similarity 22.0%; Pred. No. 2.4e-50;
 Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQGAVSGIQPLNEGVARCNESQGDVATCSCQDCAACAPALARPQALDSTFVLGMPGS 284
 Db 24 PGRPAGGRRRTTGLRRAAPDRD-----YLRPSYCDAAALSGI--- 65
 QY 285 LVLIILCSVFAVWTVLLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
 Db 66 -----SKGRATGRKAPLWLRKAFORLLFKLGCYIQNC 98
 QY 338 QGFFQGWGTWASWPLTILVSVIPVALAAGLVFTELTDPVELWSAPNSQARSEKAPH 397
 Db 99 KGF-----LVVGLLTFGAPVGLKAANTLNVELNVELVEVGGVRSBELNYT 143
 QY 398 DQHGFPFRTN-QVILTAPNRSSVRYDSLLGPKNFSGILDLDLLELE--LQERLRL 454
 Db 144 RQKIGERAMFAPQLMIQTPEEG-----ANVLTTEALLQHLDSALQASRVHV 190
 QY 455 QWSPQAQRNTSLQDICYAP-----LNPDNSTLYDCCINSLLQYFQNNRTLLLLTA 505
 Db 191 YMYN---ROWKLEHLCKYSGELITETGYMDQIBYLYPCLITPLDCFWEGAKLQSGTA 246
 QY 506 NQTLMGQ-----TSQVDKDHFLY-----CAN----- 527
 Db 247 ---YLLGKPLRWTFNPDPLFLEELKKINYQVDSWEEMLNKAFAVGHGYMDRPLCPINPADPC 304
 QY 528 -----APL-----TFKDG-----ALASGWA 544
 Db 305 PATAPNKNSTPLDMALVNLGCGHLSRKYMHQBELLVGGTVKVNSTKLSAHLQTMF 364
 QY 545 DYGAFFVFPFLAIGGYKGDKYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFQ 604

Db 191 YMYN-----RQWLEHLCKYSGBLITETGYMDQIIIEVYPCLIITPLDCFWGAKLQSGTA 246
QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFNDFLEBELKKNYQVDSWEEMLNKAEGVGHGYMDRCPCLNPADPDC 304
QY 528 -----APL-----TFKQGT-----ALALSQWA 544
Db 305 PATAPNKNSTKPLDVALVNGCHGLSRKYMHWBELIVGGTVKNSGKLVSAAHALQTMP 364
QY 545 DYCAPVPFPLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAEFLBEMRAFO 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE-----DKAAALIEAMQRTVVEVHQS 411
QY 605 RMAGHQVQVTFPAERLEDEINRTAEDLPFATSYIVIFLYISLALGSYSWSRWMD 664
Db 412 AQNSTQKVLSTFT--TTLDDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
QY 665 KATLGLGGVAVVLGVAAMGFFSYLGRSSLVILQVPPFLVLSVGAUNIPFVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLAHAFSET 525
QY 725 PRPGEPREHVHIGRALGRVAPGMLCSLSEACFFLGLALTMPAVRTEALTSLGLAVILDF 784
Db 526 GQNKRIPEDRTEGCLKRTGASVALTSISNVTAFMAALIPIPALRAFSLQAAVVVVVFN 585
QY 785 LQMSAFVALLSLDKROEASRLDVCCC-----VKPQEL-----PPPG 822
Db 586 AMVLLIFPALLSMDLYRREDRLDFCCFTSPCVSRVIOVEPOAYTDHNTRYSPPPPY 645
QY 823 QEGEL----- 827
Db 646 SSHSPAHEITQIMQSTVOLRTYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGFQKAYAPFLLHWTIRGVVLLIFLALFGVSL 861
Db 706 SSTRDLLSQFSSLSHCHLEPPCTKWLTSSFAEKHYAPFLPKAKVWIFLPLGLGLVSL 765
QY 862 YSMCHISVGLDQELAPKDSYLLDFLNFYFVFGAPVYFVTLTYGNFSSAGHNAICS 921
Db 766 YGTTVRDGLDITDVPRETRYDFTIAQKFYF-----YNN----- 803
QY 922 SAGCNPFSTQKIQVAT-----BPEGOSYLAIPASS-----WYDDFDIML----- 961
Db 804 -----YIVTQADYFNQIHLLYDLRFSFNKVKYVMLEENKQLPKWHLHYFRDWLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDFECSTVNSLCLNKMISITWGS 1000
Db 858 APDSWETGKIMPNNYKNGSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
QY 1001 VRPSVEQFHKYL--PMFLND-----RPNKCPKGLAAY--STSVNLTSDGQ 1043
Db 914 INPSA--FYIYLTAWVNDPVAYAASQANIRPHREWHDK--ADYMPETRLRIPAAB 968
QY 1044 VLASFMAHYHPLKNSDYTEALRAARELANITA--DLRKVPFGTDPFAFEPVPTITNVFY 1102
Db 969 IEYAQFPFYLNGLRDTSDFVEAIEKVRTICSNYTSLSGLSSYPNG-----YFP-----LFW 1018
QY 1103 EOYLTIPEGLPMLSLCVPTFAVSCLLGLDLRSLGNLLSIVMLIVDTVGFMAWDIS 1162
Db 1019 EOYGLRHWLLFTISVWLACTPLVCVFLNFWTAGII--VMVLALMTVELFQMGLIGIK 1077
QY 1163 YNAVSLINLVAVGMSVFPVSHITRPF--AISTKPTWLERAKEATISMGSAVFAVAMTN 1220
Db 1078 LSAVPVVLIIASVGIGVETVHVVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
QY 1221 LPGAIVLGLAKAQLIQIFPFRNLNLTLLGLLHGLVFLVILSYVG--PDVNPALALEQ- 1277
Db 1134 LLGLVLMAGSEBDFIVRYFFAVLAILTILVLINGLVLLPVLLSPGYPPEVSPANGNLRL 1193
QY 1278 --KRABEAVAAVMVASCNHPHSRVSTADNIYVNHSPFGSING 1317
Db 1194 PTPSPPPPPSVVRAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 34

AAB67163

ID AAB67163 standard; protein; 1447 AA.

XX AAB67163;

AC AAB67163;
DT 12-APR-2001 (first entry)

XX Human patched protein.

XX Human; patched; PTC; segment polarity; limb patterning; development;
KW hedgehog; antibody; mouse.
XX Homo sapiens.

OS US6172200-B1.

PN 09-JAN-2001.

XX 20-OCT-1997; 97US-00954668.

XX 07-OCT-1994; 94US-00319745.

PR 06-OCT-1995; 95US-00540406.

XX (STRD) UNIV LELAND S STANFORD.

PI Scott MP, Goodrich LV, Johnson RL;

XX WPI; 2001-136884/14.

DR N-PSDB; AAF32185.

XX Novel monoclonal antibody useful in diagnostic assays for detection of
presence of protein on surface of cells specifically binds to naturally
occurring patched protein, other than Drosophila patched protein.

PS Disclosure; Col 61-70; 39pp; English.

XX The present invention provides a monoclonal antibody which specifically
binds to a patched protein (PTC) other than that from Drosophila. Also
given are the protein and coding sequences of patched from the beetle,
mouse, butterfly, fruit fly, mosquito and human. Patched is a segment
polarity gene involved in limb patterning. The sequences can be used to
study development and to isolate the patched ligand, hedgehog. In
addition, antibodies can be used to detect the PTC protein on cell
surfaces or to inhibit the transduction of signal by the PTC ligand by
competing for its binding site

SQ Sequence 1447 AA;

Query Match 9.4%; Score 651.5; DB 4; Length 1447;

Best Local Similarity 22.0%; Pred. No. 2.4e-50;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGOAVSGIGIPLNGSVARCNESQGDVATCSCQDCASCPAIPALQALDSTFYLGQMPGS 284

Db 24 PGRFAGGRRRTTGLRRRAAPDRD-----YLRPSYCDAAFALEQI--- 65

QY 285 LVLIILILCSVPAVVTVLLVGFVAPARDKSRWVDPK-----KGTSLSDKLSFSTHTLL 337

Db 66 -----SKGATGKAPLWLRKAFQRLFLKLCYIQKNC 98

QY 338 GQPFQGHGTWVASWPLTILVSVIPVVALAAGLVFTELTTPDVELWMSAPNSQARSEKAPH 397

Db 99 GKF-----LVVGLLIIFGAFVGLKAAANLETNVBEVLWVEVGRVSRRLNYT 143

QY 398 DQHFQPPFRTN-QVILTAPNRSYRYDLSLLGPKNFSGILDLLELLS--LQERLRL 454

Db 144 RQKLGSEAMFPQMIQTPEEG-----ANVLTTEALLQLHLSALQASRVHV 190

QY 455 QVWSPEAQRNISLQDICVAP-----LNPDNSTSLYDCCINSLLQYFONNRTLLLLTA 505

191	Db	191	YMYN----	QWKL	EHCH	YKSG	BELIT	ETG	MDQII	EYLP	CLII	ETPL	DCFW	EGAK	LQSG	TA	246							
506	Qy	506	NOTLMO-----	TSQV	MDKH	FLY-----	CAN-----	527																
247	Db	247	--YLLG	KPPL	RWTN	FPD	PLF	BEEL	KKIN	QVDS	WEEM	LKAB	RGV	HGMD	RPC	LN	PA	304						
528	Qy	528	-----APL-----	TFK	DGT-----	ALAL	SCMA	544																
305	Db	305	PATAPN	KNI	STK	P	LD	MA	LV	NG	CH	GL	SR	KY	MH	Q	BE	1	364					
545	Qy	545	DYGA	PVP	FLA	GGY	KDY	SE	AE	AL	IM	TF	SL	NN	PAG	PR	LA	QAK	1	604				
365	Db	365	QMT	PK	QWY--	EH	PK	GYEY-----	VSH	INNE-----	DKA	AA	LE	AQ	WT	VE	VH	QSV	411					
605	Qy	605	RRMAG	MQ	VTT	PA	ERS	LE	BI	NT	TA	ED	PI	FAT	SY	IV	FL	IS	AL	1	664			
412	Db	412	AQNS	TQ	VL	SE	TT--	TL	DD	IL	K	S	F	SV	IR	VAS	G	YL	LA	1	465			
665	Qy	665	KATL	GG	VA	VL	GA	VM	AA	MP	FS	YL	GI	RS	SL	VIL	Q	VP	FL	1	724			
466	Db	466	QGA	VL	AG	VL	VAL	SV	ANG	GL	SL	GI	S	NA	AT	TQ	VL	FL	AL	1	525			
725	Qy	725	PR	PG	B	PRE	VH	I	GA	L	GR	VA	PS	ML	CS	LS	E	A	I	1	784			
526	Db	526	GQ	NK	R	I	P	E	D	T	G	E	C	L	K	R	T	G	A	S	1	585		
785	Qy	785	LLQ	S	A	F	VAL	L	S	D	S	K	Q	E	A	S	R	L	D	V	1	822		
586	Db	586	AM	VL	I	F	AIL	US	M	D	T	Y	R	R	D	I	F	C	C	T	1	645		
823	Qy	823	QEG	L-----	827																			
646	Db	646	SSH	S	PA	HE	T	I	M	Q	S	T	V	Q	L	R	E	D	P	H	1	705		
828	Qy	828	-----	LL	GF	K	A	Y	A	P	FL	L	H	W	T	R	G	V	L	L	1	861		
706	Db	706	S	S	T	R	D	L	S	Q	F	S	S	L	H	C	L	E	P	C	T	1	765	
862	Qy	862	Y	S	M	C	H	I	S	V	G	L	Q	E	A	L	P	K	D	S	Y	1	921	
766	Db	766	Y	T	T	R	V	R	G	L	D	T	I	V	P	R	E	T	R	E	D	1	803	
922	Qy	922	S	A	G	C	N	F	S	T	Q	I	Q	V	A	T-----	E	P	Q	S	Y	1	961	
804	Db	804	-----	YI	T	K	A	D	Y	P	N	I	Q	H	L	L	H	R	S	F	N	1	857	
962	Qy	962	-----	TP	S-----	C	R	L	Y	I	S	G	N	K	D	K	F	C	P	T	V	1	1000	
858	Db	858	A	F	D	S	W	E	T	G	K	M	P	N	Y	K	N	G	S	D	D	1	913	
1001	Qy	1001	V	R	P	S	V	Q	E	F	H	K	Y	L-----	P	N	I	K	P	K	G	1	1043	
914	Db	914	I	N	P	S	A-----	F	Y	I	L	T	A	M	S	N	D	P	V	A	A	1	968	
1044	Qy	1044	V	L	A	S	R	M	A	H	K	L	K	N	S	O	D	T	A	L	R	1	1102	
969	Db	969	I	B	A	P	F	F	Y	L	G	R	D	T	S	F	V	E	A	E	K	1	1018	
1103	Qy	1103	E	Q	V	L	I	P	B	G	L	E	M	L	S	I	C	L	V	P	A	S	1	1162
1019	Db	1019	E	Q	I	G	L	R	H	W	L	L	F	I	S	V	L	A	C	T	F	1	1077	
1163	Qy	1163	Y	N	A	V	S	L	I	N	V	S	A	G	H	S	V	E	F	V	S	H	1	1220
1073	Db	1073	L	S	A	V	P	V	I	L	A	S	V	G	I	E	V	T	H	A	F	A	1	1133
1221	Qy	1221	L	P	C	I	L	V	I	G	L	A	Q	A	I	O	I	P	F	F	R	1	1277	
1134	Db	1134	L	L	G	V	L	M	A	G	S	E	D	F	I	V	R	E	F	A	V	A	1	1193
1278	Qy	1278	--	K	R	A	E	A	V	A	A	V	M	A	S	C	N	P	H	R	S	1	1317	
1194	Db	1194	P	T	S	P	P	P	P	S	V	R	F	A	M	P	P	G	H	T	H--	1	1233	

RESULT_35	
AAE19830	
ID	AAE19830 standard; protein; 1447 AA.
AC	AAE19830;
XX	
XX	18-JUN-2002 (first entry)
DT	
XX	
XX	Human patched (Ptch) protein.
DE	
XX	Human; patched; Ptch; cell proliferation; differentiation; therapy;
KW	cystostatic; testicular cancer; hedgehog protein signalling.
KW	
OS	Homo sapiens.
OS	
XX	
FH	Location/Qualifiers
FT	Domain
FT	/label= TM1
FT	/note= "Transmembrane domain"
FT	438..458
FT	/label= TM2
FT	/note= "Transmembrane domain"
FT	471..493
FT	/label= TM3
FT	/note= "Transmembrane domain"
FT	502..521
FT	/label= TM4
FT	/note= "Transmembrane domain"
FT	548..570
FT	/label= TM5
FT	/note= "Transmembrane domain"
FT	578..599
FT	/label= TM6
FT	/note= "Transmembrane domain"
FT	751..769
FT	/label= TM7
FT	/note= "Transmembrane domain"
FT	1028..1048
FT	/label= TM8
FT	/note= "Transmembrane domain"
FT	1052..1074
FT	/label= TM9
FT	/note= "Transmembrane domain"
FT	1078..1097
FT	/label= TM10
FT	/note= "Transmembrane domain"
FT	1122..1141
FT	/label= TM11
FT	/note= "Transmembrane domain"
FT	1155..1178
FT	/label= TM12
FT	/note= "Transmembrane domain"
XX	
PN	US6348575-B1.
XX	
PD	19-FEB-2002..
XX	
PF	15-APR-1999; 99US-00293505.
XX	
PR	15-APR-1999; 98US-0081884P.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	De Sauvage F, Carpenter DA;
XX	
DR	WPI; 2002-215260/27.
XX	
PT	Native human patched-2 polypeptide for treating disorders caused by
PT	Hedgehog protein signaling such as testicular cancer, and for screening
PT	cdNA libraries.
XX	

Example 1; Fig 3; 82pp; English.

PS The invention relates to an isolated sequence comprising a native human
 CC patched2 (Ptc2) polypeptide. The invention also relates to signalling
 CC and mediator molecules in the hedgehog (hh) cascade which are involved in
 CC cell proliferation and differentiation. The isolated sequence is useful
 CC for the treatment of disorders which are linked to Hedgehog, especially
 CC used as a hybridisation probe in a cDNA library to isolate Ptc2 or its
 CC homologues, and to diagnose whether a disorder is driven by Ptc2-2 or
 CC Hedgehog protein signalling. The present sequence is human patched (Ptc2)
 CC protein. Note: The sequence shown as SEQ ID NO: 4 in sequence listing of
 CC fig 3 of the specification
 CC
 XX

Sequence 1447 AA;

Query Match 9.4%; Score 651.5; DB 5; Length 1447;

Best Local Similarity 22.0%; Pred. No. 2.4e-50;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGOAVSGGIQPLNEGVARCNESQDDVATCSQDCAASCPAIAIPQALDSTFVLGMPGS 284

DB 24 PGPAGGRRRTGRLRAAAPDRD-----YLRPSYCDAAALEQI--- 65

QY 285 LVLIILCSFVAVVITLLVGFVRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337

DB 66 -----SKGKATGRKAPLWLRAKQFRLFLKGCYIQKNC 98

QY 338 GQPFQGWGTWASWPLTILVSVIPVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFH 397

DB 99 GRP-----LVGGLIFGAPAVGLKAANLTNEVELWVEVGRVSRLENTY 143

QY 398 DQHFQPPFRPTN-QVILTAPNRSSRYDLSLLGPKNFGSLDLDLLELE--LQERLRL 454

DB 144 RQKIGEEAMFNQLMITQPKREG-----ANVLTEALLQHLDSALQASRVHV 190

QY 455 QVMSPEAQARNISLQIDICYAP-----LNPDTSLYDCCINSLLIQYQNNRTLLILTA 505

DB 191 YMYN---RQWLEHLCKYSGBELITETGYMDQIIEVLYPCLITPLDCFWGAKLQSGTA 246

QY 506 NOTLMGQ-----TSQVDMKQHFY-----CAN----- 527

DB 247 ---YLLGKPLRWNTFDPLFELBELKKINYQVDSWEMLNKAEBVGHGYMDRPLNADPDC 304

QY 528 -----APL-----TFKQGT-----ALALSQWA 544

DB 305 PATAPNKNSTKPLDMLVNGCHGLSRKYMWQBELIVGGTVKNSGKLVSAHALQTMF 364

QY 545 DYCAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKMBEAFLEENRAPQ 604

DB 365 QLMTPKQMY---EHPKGYEY-----VSHINWNE---DKAAALLENMQRTYVEVHQSV 411

QY 605 RMAGMQVQVTPAERSLDEINRTTADLPPIFATSYIVIFLYISIALGYSYNSWRVMVDS 664

DB 412 AQNSTOKVLSFTT-ITLDDILKSFSDSVIRVASGYLLMLAVACLTM---LRWD---CSKS 465

QY 665 KATLGGVAVVLGAVMAWGFYSYLGIRSSLVILQVPPVPLVLSVCAADNIFIPVLEYQRL 724

DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNATTTQVLPPLGALGVGDVDFVLAHAFSET 525

QY 725 PRPQGPREVHIGRALGRVAPSMILCSLEALCFPLGALTMPAVRTFALTSLGLAVILDF 784

DB 526 GQNKRIFFEDRTGECLKRTGASVALTSISNTVAFPMALIPALPAFSLQAQVAVVNF 585

QY 785 LLOMSAFVALLSDSKRQASRLDVCCC-----VKQDEL-----PPPG 822

DB 586 AMVLLIFPAILSMDLVRRDRDLDFCCPTSPCVSRVIOVEFOAVTDTHTNTRYSPPPY 645

QY 823 QDEGL----- 827

DB 646 SSHSAFHEQTITMQSTVQLRTEYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSCOSPEST 705

QY 828 -----LLGFFQKAYAPFLLHWTITRGVLLFLALFGVSL 861

DB 706 SSTDRLLSQFSSSLHCLPEPPCTKWTLSFAEKHYAPFLLKPKAKVWVIFLGLLGVSL 765

QY 862 YSMCHISVGLDDELALPKDSYLLDFPLNRYFVEGAPVYFTTILGYNSSSRGAMNACS 921

DB 766 YGTTVRDGLDITDVPRETREYDFAAQKYPSP-----YNN----- 803

QY 922 SAGCNFFSTQKIQYAT-----EPQESYLAIPASS-----WVDFIDWL----- 961

DB 804 -----YIVTQADYPTNQHLYDLHRSFNVKYVMLEENKQLPKMWLHYFRDQLQLOD 857

QY 962 -----TPSS-----CCRLYISGPNKQKFCPSTVSNLCLNKMCSITMGS 1000

DB 858 AFDSDWETGKIMPNNYKNGSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913

QY 1001 VPSPVEQHKYL-PWFLND-----RPNKCPKGLLAAY--STSVNLTSQ 1043

DB 914 INPSA--FYIYLTAWVSNDFVAYAAASQANIRPHRPFWHDK---ADYMPETRLRIPAAEP 968

QY 1044 VLASRFMAYHKPKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPFVFPVTTINVFY 1102

DB 969 IEYAQPPFVYNGRLDTSDFEAEIKVRTICSNYTSGLSSYPNG-----YDF-----LFW 1018

QY 1103 EQYLTPLEGLPMLSLCVPTFAVSLGLDLRLSLMLLSIVMLVDTVGFMAWDLS 1162

DB 1019 EQYIGLRHWLLLFISWLACTFLVCAPFLNPNWTAGII-VMVLAALMTVELFCGMGLIGIK 1077

QY 1163 YNAVSLINLVSAVMSVRFVSHITSP--AISTKPTWLERAKENTISMGSAVPAGVAMTN 1220

DB 1078 LSAPVWVILIASVGVIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133

QY 1221 LPGIILVLGLAKAQLIQIPFRLNLITLIGLHGLVFLPVLISYVG--PDVNPALALEQ- 1277

DB 1134 LGVLMLAGSEDFIVRFFAVLAITITLGLVNLGLVLLPVLISFGFPGYPEVSPANGNLRL 1193

QY 1278 --KRAEEVAAMVASCPNHPNSRVSTADNIYVNHSPFSGSIK 1317

DB 1194 PTPSPPPSVVRFAAMPQHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 36

ABJ10931

ID ABJ10931 standard; protein; 1447 AA.

XX AC ABJ10931;

XX DT 12-DEC-2002 (first entry)

XX TRC8 related human patched protein SEQ ID No 5.

XX TRC8; Translocation in Renal cancer from Chromosome 8; fused DNA; 3.2;

XX KW FHIT/TRC8 fusion DNA; sporadic renal cell carcinoma; TRC8/FHIT; TRC8FHIT;

XX KW human chromosomal translocation.

XX OS Homo sapiens.

XX PN US2002106656-A1.

XX PD 08-AUG-2002.

XX PF 02-JUL-2001; 2001US-00898533.

XX PR 12-MAR-1998; 98US-0077723P.

XX PR 12-MAR-1999; 99US-00268140.

XX XX (GEMM/) GEMMILL R M.

XX PA (DRAB/) DRABKIN H A.

XX PI Gemmill RM, Drabkin HA;

XX WPI; 2002-712395/77.

XX

[illegible]

XX Human patched gene PTC product.
DE Human; patched gene; PTC; hedgehog protein; gene therapy.
KW Homo sapiens.
XX OS US2003032085-A1.
XX PN 13-FEB-2003.
XX PD 20-OCT-1997; 97US-00954701.
XX PF 07-OCT-1994; 94US-00319745.
XX PR 06-OCT-1995; 95US-00540406.
XX PA (SCOTT/) SCOTT M P.
PA (GOOD/) GOODRICH L V.
PA (JOHN/) JOHNSON R L.
XX Scott MP, Goodrich LV, Johnson RL;
XX WPI; 2003-492065/46.
XX New DNA sequence other than present in a chromosome encoding patched gene
PT other than Drosophila patched gene, useful for preparing transgenic
PT laboratory animals and to knock out patched protein in embryonic stem
PT cells.
XX Disclosure; Page 8-10; 40pp; English.
XX The invention relates to a DNA sequence other than present in chromosome
CC encoding a patched (ptc) gene other than Drosophila patched gene. A cell
CC expressing the DNA sequence is useful for producing patched protein, by
CC growing the cell expressing the DNA sequence, where the patched protein
CC is expressed and isolating the patched protein free of other proteins.
CC The cell expressing the DNA sequence is also useful for screening
CC candidate compounds for binding affinity to the patched protein, by
CC combining the candidate protein with the cell expressing the DNA sequence
CC where the DNA sequence comprises the entire coding sequence under the
CC transcriptional regulation of the transcriptional initiation region and a
CC transcriptional regulation region functional in the cell, expressing the
CC patched protein in the cell and assaying for the binding of the candidate
CC compound to the patched protein. The above cell is useful for screening
CC candidate compounds for agonist activity with the patched protein. The
CC DNA sequence is useful for producing all or portions of the patched
CC protein, as probes for research, diagnosis, binding of hedgehog protein
CC for its isolation and purification and in gene therapy. The DNA sequence
CC is also useful as primers for investigating other species and for
CC isolating genes from various mammalian sources of interest, particularly
CC from humans or from domestic animals. The DNA sequence is further useful
CC for preparing transgenic laboratory animals and to knock out the PTC
CC protein in the embryonic stem cells, so as to produce hosts with single
CC functional patched gene. The present sequence represents the amino acid
CC sequence of the human patched gene PTC product
XX
SQ Sequence 1447 AA;
Query Match 9.4%; Score 651.5; DB 7; Length 1447;
Best Local Similarity 22.0%; Pred. No. 2.4e-50;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY 225 PQAVSGGIQPLNEGVARCNESQDDVATSCQDCAACPAIARPOALDSTFVLGMPGS 284
DB 24 PGRPAGGRRRTTGLRAAAPDRD-----YLRPSYCDNAFALEI--- 65
QY 285 LVLIILCSFVAVVTILLVGRVAPARDKSKWDPK-----KGTSLSDKLSFSFTHLL 337
DB 66 -----SKGATGRKAPLWLRKAFQRLFLKGCYIQKNC 98
QY 338 GQFPQGWGTVWASWPLTILVSVIPVVALAAGLVFTLTDTPVELLSAPNSQARSEKAPH 397
DB 99 GK-----LVVGLLIIFGAFVAGLKAANLETNVEELWVEGVGRVSRRLNYT 143

QY 398 DOHFGPPFRTN-QVILTAPNRSSRYDSSLGLGPKNFSGILDLLELLELLE--LQERLRHL 454
DB 144 RQKIGSEAMFNPQIMIQTPKEBG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QVMSPEAQRNISLQDICYAP-----LNPDNTSLYDCCINSLSLQYFQNNRTLLLLTA 505
DB 191 YMYN-----RQWLEHLCKYKSGELITETGYMDQIIIEVLYPCLIITPLDCFWEGAKLSGTA 246
QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
DB 247 --YLLGKPLRLWTNFDPLELEELKKINQVDSWEEMLNKAEVGHGWMDRPCLNPADDDC 304
QY 528 -----APL-----TPKQGT-----ALALSCMA 544
DB 305 PATAPKNKSTKPLDMALVNLGGCHGLSRKYMHWQEBELIVGTVKNSGKLVSAAHALQTMF 364
QY 545 DYGAIPVFPFLAIGGYKGYKDYSEABALIMTSLANNYPAGDPRLAQAQAKLWEBAFLMEARAFQ 604
DB 365 QLMTPKQMY--BHFPGYGY-----VSHINWNE---DKAAAILLEAWQRTYVEVHVQSV 411
QY 605 RRMAGMEQVTFPFAERSLEDEINRTAEDLPIFATSYIVIFLYISLALGSSYSSSRVMVDS 664
DB 412 AQNSTQKVLSTFT--TTLDDILKSFSDVSIVRVSAGYLLMLAYACLTH---LRWD--CSKS 465
QY 665 KATLGLGSAVAVLGAVMAAMGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLALGVGVDDVFLLAHAFSET 525
QY 725 PRPGEPREVHIGRALGRVAPSMMLLCSLSEAI CFFGALTPMPAVRFTALTSGLAVILDF 784
DB 526 GONKRIPEFDRGTGCLKRTGASVALTSISNVTAFMAALIPALPAFSLQAQVVVVFNF 585
QY 785 LLQMSAFVALLSLDSKQESRLDVCC-----VKQBEL-----PPPG 822
DB 586 AMVLLIPAILSMOLYRREDRRLDIFCCFTSPCVSRVIOQVEPQAYTHTDNTRYSPPPY 645
QY 823 QGEG----- 827
DB 646 SSHFAHETQITMSTVOLREYDPTHVYTTAEPRSEISQPVVTVTQDTLSCQSPEST 705
QY 828 -----LLGFPKAYAPFLHWHITRGVLLFLFLALFGVSL 861
DB 706 SSTRLDLSQFSDSSHLCHLEPCTKWTLSFSAEKHYAPFLKPKAKVQVIVIFLGLLGVSL 765
QY 862 YSMCHISVGLDQELAPKDSYLLDYFLNRYFEVGAPEVFTVLGYNFSSEAGMNAICS 921
DB 766 YGTRVRDGLDLTDIVPRETREYDFIAAQKYFSP-----YNN----- 803
QY 922 SAGCNMFSTQKIQYAT-----BFPQSYLAIPASS-----WVDDFDWL----- 961
DB 804 -----YIVTQKADYPNIQHLLYDLHRSFNKVMYLEENKQLPKMWLHYFRDWLQGIQD 857
QY 962 -----TPSS-----CCRLYISGPNKDKFCPSTVNSLCLNCKMSITMGS 1000
DB 858 AFDSDMETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDAQGI 913
QY 1001 VRPSVEQPHKYL-PWFLND-----RENKCPKGLAAY--STSUNLTSQDQ 1043
DB 914 INPSA--FYIYLTAWVSNDPVAYAASQANIRHPREWVHDK---ADYMPETRLRIPAAEP 968
QY 1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA-DLRKVPCTDPAPAEVPPYTTINVPY 1102
DB 969 IEYAGFPFPLNGLRDTSDFEAIKERTICSNTYSLGLSSYPNG-----YFF-----LFW 1018
QY 1103 EOYLITLPEGLPMLSLCLVPTFAVSCLLGLDLRSLNLLSIVMILVDVTGVMALWDIFS 1162
DB 1019 EOYIGLRHWLLFISVLACTIONFLVCAVFLNFWTAGIIVMVLALMTVELFCWMGLIGK 1077
QY 1163 YNAVSLINIVSAGVSVFVSHITRSP--AISTKPTWLERAKEATISMGSAVAGVAMTN 1220
DB 1078 LSAVPVILIASVGIGVBTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133

QY 1221 LPGILVLGLAKAQLIOIFPPRLNLLTLGLLGLVFLPVLVYG--PDVNPALALEQ- 1277
Db 1134 LLGVLMLAGSEDFIVRYFFAVLAILTLGLVLLPVLFFGPFYEVSPANGLNRL 1193
QY 1278 --KRAEEAANVAVSCNHPHSRVSTADNIYVNHSPGSGIKG 1317
Db 1194 PTPSPPPPPSVRFAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 39

ADD46678
ID ADD46678 standard; protein; 1447 AA.
XX ADD46678;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Human Protein NP_000255, SEQ ID NO 12363.
XX

KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX

OS Homo sapiens.

OS Unidentified.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GENO) GEN HOSPITAL CORP.

PA (PABB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; NP_000255.

XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Example 1; Page; 1017pp; English.
XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3

CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1447 AA;

Query Match 9.4%; Score 651.5; DB 7; Length 1447;
Best Local Similarity 22.0%; Pred. No. 2.4e-50;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQAVGSGIQPLNEGVARCNESQGDVATCSCDCAACPATARPQALSTFVLGMPGS 284
Db 24 PCRAGGGRRRRTGGLRAAAPDRD-----YLRPSYCDAAFALEQI--- 65

QY 285 LVLIILICSVPAVAVTILLVGPVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWLRKAFORLLFKLGCYQKNC 98

QY 338 GQFFQGWGTWVASWPLTILVSVIPVVALAAGLVFTTDDPVELWSAPNSQARSEKAFH 337
Db 99 GKF-----LVGGLIFGAFVGLKAAANLSTNVEELWVEVGGRVSRRLNT 143

QY 398 DQHPGPPFTN-QVILTAPNRSRYDLSLLGPKNPSGILDLDLLELE--LQERLRLH 454
Db 144 RQKIGEEAMFNPQLMIQTPKBERG-----ANVLTTTEALIQHLDLSALQASRVHV 190

QY 455 QVWSPEAQRNISIQCICVAP-----LNPDNTSLYDCCINSLLQYQNNRTLLLLTA 505
Db 191 YMN-----RQWKLEHLCKYKSGBLITETGYMDQIIEYLYPCLITPLDCFWGAKLQSGTA 246

QY 506 NOTLMGQ-----TSQVDWKOHFLY-----CAN----- 527
Db 247 --YLLGKPLRWTFNPDPLEFLELKKINQVDSWEMLNKAEVGHGYMDRCPCLNPADPC 304

QY 528 -----APL-----TFKQGT-----ALALSCWA 544
Db 305 PATAPNQNSTKPLDMALVINGGCHGLSRKYMHWQBELIVGGTVKNSLGLVSAHALQTMF 364

QY 545 DYGAVPFPPLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLEEMRAFQ 604
Db 365 QLMTPKQMY---EHPKGYE-----VSHINNE-----DKAAALEAWQRTVVVVHQSV 411

QY 605 REMAGMFVFTTAEBSLEDEINRTTAEOLPIPATSYIVIFLYISLALGSYSSWSRVWVDS 664
Db 412 AQNSTQKVLSTFT--TLDLILKSPDSVIRVASGYLLMLAYACLTM--LRWD--CSKS 465

QY 665 KATLGLGGVAVVGLVAMAMGPFYSYLGRSSVLIVQVFPFLVLSVGADNIFIPVLEQR 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVGVDVFLAHAFSET 525

QY 725 PRPGEPREHVHIGRALGRVAPSMILCSLSEALCFELGALTMPMPVARTFALTSLGLVILDP 784
Db 526 GQNKRIPEDRTEGCEKRTGASVALTSISNVTAFPMALIPIPALRAFSLQAAVVVVFNF 585

QY 785 LIQMSAFVALLSLDSKROEASRLDVCCC-----VKPQEL-----PPRG 822
Db 586 AMVLIFPALLSMDLYREDRLDIFCCFTSPCVSRVIOVEFOAYTDTHDTRYSPPPY 645

QY 823 QGEGE----- 827
Db 646 SSHSPAHEQTITMQSTVQLRTYEDPHTVYVYTAEPGRSEISVQPVTVTQDTLSCQSPST 705

QY 828 -----LLGFPKAYAPFLLHWITRGVLLLLFALFGVSL 861
Db 706 SSTRDLLSQFSDSSLHLEPPCTKWTLSFPAEKHYAPFLPKPAKVWIFLGLGLVSL 765

QY 862 YSMCHISVGLDQELAPKDSYLLDYFLNRYFEVGAPVYFVTTILGYNFSRAGMNAICS 921
Db 766 YGTTVRGLDLDTDVPRETREYDFIAAQFKIFSP-----YMN----- 803

QY 922 SAGCNPFSTQKIQVAT-----BFPQSYLAIPASS-----WDDDFIDWL----- 961

Db 804 -----YIVTQADYVNTQHLLYDLHRSFSNVKVMLEENKQLPKMMLHYFRDWLQGLQD 857
Qy 962 -----TPSS-----CCRLYISGNKDKFCPSVTNSLNCNKMISITWGS 1000
Db 858 AFDSWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQPHKYL--PWFLND-----RPNIKCPGGLAAY--STSVMLTSDGQ 1043
Db 914 INPSA--FVIYLTAMVSDNPVAYASQANIRPHRPEWVHDK---ADYMPETLRIPAAEP 968
Qy 1044 VLASFMAHFKLNSQDYTEALRAARELAANITA-DLRKVPGTDPAPVEPPYITITNVFY 1102
Db 969 IEYAFPPFVNLGRDTSDFVEAIEKVRTICSNYISLGLSSYPNG-----YFP-----LFW 1018
Qy 1103 EOYLILPEGLPMLSLICLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGVFMALWDIS 1162
Db 1019 EQYIGLRHWLLLFISVLACTFLVCVFLNPNWTAGII--VMVLMALTVELFGMGLIGIK 1077
Qy 1163 YNAVSLINLVASGVSEFVSHITRSP--AISTKPTWLERAKEATISGSAVFAGVAMTN 1220
Db 1078 LSAVPVILIASVGIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
Qy 1221 LGCILVLGLAKAQLIQIFFRNLALITLGLLHGLVELPVLVSYVG--PDVNPALALBQ- 1277
Db 1134 LIGVLMLAGSEFDFVRYFFAVLAILITLIGVLNGLVLLPVLVLSFGFPGPYEVPSPANGIARL 1193
Qy 1278 --KRAEEAVALVAVSCPNHPSRVSVDADNIYVNHSPFGSISK 1317
Db 1194 PTPSPPPSVVRFPMPGHTH--SGSDSSDSEYSSQTTVSG 1233

RESULT 40
ADE94224
ID ADE94224 standard; protein; 1447 AA.
XX ADE94224;
DT 12-FEB-2004 (first entry)
XX Human ptc polypeptide.
XX Human; patched gene; ptc; developmental abnormality;
XX proliferative disorder; tumour; oncogenic patched mutation;
XX patched-associated phenotype; basal cell nevus syndrome; BCNS; carcinoma;
XX meningioma; fibroma; cancer.
XX Homo sapiens.
XX US6551782-B1.
XX 22-APR-2003.
XX 28-NOV-2000; 2000US-00724631.
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX 31-MAY-1996; 96US-00656055.
XX 22-AUG-1997; 97US-00918658.
XX (STRD) UNIV STANFORD.
XX (REGC) UNIV CALIFORNIA.
XX Scott MP, Goodrich LV, Johnson RL, Epstein E;
XX WPI; 2003-615308/58.
XX N-PSDB; ADE94223.
XX Assay for phenotyping patched status of cell in mammalian (preferably
XX human) cell sample, involves detecting presence or absence of aberrant
XX modification or mutation of a patched gene, and mis-expression of patched
XX gene.
XX Disclosure; SEQ ID NO 19; 57pp; English.

XX The invention relates to an assay for phenotyping the patched status of a
CC cell, involving detecting in vitro, in a sample of mammalian cells, the
CC presence or absence of a genetic lesion in a patched gene (ptc)
CC characterised by at least one of an aberrant modification or mutation of
CC a patched gene or mis-expression of the patched gene. The invention also
CC relates to a method for diagnosing a genetic predisposition of an animal
CC to a developmental abnormality or a proliferative disorder, where the
CC developmental abnormality or proliferative disorder correlates with
CC aberrant expression or activity of a patched gene or gene product.
CC involving detecting in vitro the presence of a predisposing mutation in a
CC patched gene in cells of the animal. Characterising the phenotype of a
CC tumour, involves detecting the presence of an oncogenic patched mutation
CC in cells of the tumour, where the presence of the oncogenic mutation
CC indicates that the tumour has a patched-associated phenotype. The assay
CC is useful for phenotyping the patched status of the cell in a mammalian
CC cell sample obtained from a human patient. The disorders that can be
CC detected include basal cell nevus syndrome (BCNS), carcinomas,
CC meningiomas and fibromas. This sequence represents a human ptc
CC polypeptide of the invention.
XX
XX Sequence 1447 AA;
SQ

Query Match 9.4%; Score 651.5; DB 7; Length 1447;
Best Local Similarity 22.0%; Pred. No. 2.4e-50;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PGQAVGSGIQPLNKGVARCHESQDDVATCSCQDCQACPAIARPOALDSTFYLGQMPGS 284
Db 24 PGRPAGGRRRTTGLRRAAPDRD-----YLRPSYCDAAFALEQI--- 65

Qy 285 LVLIILCSVFAVVVITLLVGRVAPARDKSKWDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGRATGKAPLWLRKAFQRLFLKLCGYIQKNC 98

Qy 338 GQFFQMGCTWASWPLTILVLSVIPVVALAGLVFTTELTDPVELWSAPNSQASEKAFH 397
Db 99 GKF-----LVVGLLIFGAFVGLKAAANLETVNEELWVEVGVGRVSRRELYT 143

Qy 398 DQHFGPPFRTN-QVILTAPNRSYRYDSLLGPKNFSGIILDLLELE--LQRLRLH 454
Db 144 RQKIGEEAMFNPQLMIQPKSEG-----ANVLTEALLQHLDSALQASRVHV 190

Qy 455 QVMSPEAORNTSLQICYP-----LNPNTSLYDCCINSLLQVFNRTLLLTATA 505
Db 191 YMN---RQWKLHLCYKSGELITETGYMDQIIEYLPCLITPLDCFWEGAKLQSGTA 246

Qy 506 NQTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPPLRWTFNDFPLEFLEELKKNYQVDSWEEMLNKAEVGHGYMDRPCINPADPDC 304

Qy 528 -----APL-----TFKDG-----ALALSCMA 544
Db 305 PATAPNKNSTKPLDMALVNGCGHLSRKYMWQBELVGGTVKNTGKLVSAHALQTMF 364

Qy 545 DYCAPVFPFLATGKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEENRAFQ 604
Db 365 QLMTPKQWY---EHFGYGYEY-----VSHINWNE---DKAAAILAEWQRTYVEVVHQSV 411

Qy 605 RRMAGHFQVTFPAERSLEDEINRTAEDLPITATSYIVIFLYISLALGSYSSRWVDS 664
Db 412 AQNSTQKVLSTFT--TTLDLILKFSQVSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465

Qy 665 KATLGLGVAVVLGAVMAAMGFFSYLGRSSIVLQVWPFVLVSUGADNIFIFVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNATTOVLPFLALGVGVDDVFLLAHAFSET 525

Qy 725 PRPGEPREVHI GRALGRVAPSNMLCSLSEACFFFLGALTTPMPAVTFTALTSLAILDF 784
Db 526 GQNKRIFFEDRTGECIKRTGASVALTSISNVTAFMAALIPALRAFSLQAAVVVVFNF 585

Qy 785 LQMSAFVALLSLDSKQESRLDVCC-----VKQEL-----PPPG 822

586	AMVLLIPEAILSMYLRREDRLDIFCCFTSPCVSRVQVEPOAYTDTHTDNTRYSPPPY	645
823	QGEGL-----	827
646	SSHSFAHETQITMSTVOLRTEYDPHTVYTTAEPRESEISVQPVTVTQDTLSCQSPST	705
828	-----LLGFFQKAYAPFLHLWITRGVLLFLFALFGVSL	861
706	SSTRDLSQFSDSLHCLPEPCTKWTLSGPAEKHYAPFLKPKAKVVVIFLGLLGVSL	765
862	YSMCHISVGLDQELAKPDSVLLDYFLFNRYFEVGPVVFVTVTLGVNFSSEAGMNAICS	921
766	YGTRVRDGLDITVPRETREYDFIAAQKPYFSF-----YNM-----	803
922	SAGCNFSFTKIOVAT-----EFPEQSQSLAIPASS-----WVDDPFIDWL-----	961
804	-----YIVTKADYPNIOHLLYDLHRSFSNVKVMLEENKQLPKMHLHYFRDMLQGLQD	857
962	-----TPSS-----CCRLYISGPNKOKFCSTVNSLNLCKNCHSITWGS	1000
858	AFDSWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI	913
1001	VRPSVEQHKYL-PWFLND-----RPNIKCPKGGAAV--STSVNLTSQGQ	1043
914	INPSA--FYIYLTAWVSDNPVAYAAQANIRPHRPEWVHDK---ADYMPETRLRIPAAEP	968
1044	VLASRFMAYHKPKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPFVPPYITNVFY	1102
969	TEYAQFPFVGLRLRDTDFVEAIEKVRTICSNYTSGLSSYPNG-----YPF-----LFW	1018
1103	EQYLTLPEGLPMLSLCIVPTFAVSCLLGLDLRSGLNLLSIVMLYDVTGFWALWDIS	1162
1019	EQYIGLRHLLLFISVLACTFLVCVFLNFWTAGII-VMLALMTVELFGMGLIGIK	1077
1163	YNVSLINLVASGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSVAFGAVAMTN	1220
1078	LSAVPVVILLIASVGIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST	1133
1221	LPQILVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPVILSVYG--PDVNPALALEQ-	1277
1134	LLGVLMLAGEPDFVIRYFFAVLAILITLGLVNLGLVLLPVLLSFFGPPYEPSPANGLNRL	1193
1278	--KRAEEAVAAVWASCPNHPRSVSTADNIYVNHGPEGSIKG	1317
1194	PTPSPRPSPSVVRFAFPFGHTH--SGSDSDSEYSSQTTVSG	1233
RESULT 42		
DB	AD48989	
XX	AD48989 standard; protein; 1447 AA.	
XX	AD48989;	
XX	29-JAN-2004 (first entry)	
XX	Human patched protein (ptc).	
XX	cytostatic; vulnery; gene therapy; phenotyping; patched status;	
XX	patched gene; genetic predisposition; basal cell nevus syndrome; tumour;	
XX	carcinoma; meningioma; medulloma; fibroma; cancer; wound healing; aging;	
XX	human; patched gene; ptc.	
XX	Homo sapiens.	
XX	.US2003186309-A1.	
XX	02-OCT-2003.	
XX	22-APR-2003; 2003US-00421446.	
XX	07-OCT-1994; 94US-00319745.	
XX	06-OCT-1995; 95US-00540406.	
XX	31-MAY-1996; 96US-00656055.	

PR	22-AUG-1997;	97US-00918658.				
PR	28-NOV-2000;	2000US-00724631.				
PA	(STRD)	UNIV LELAND STANFORD JUNIOR.				
PA	(REGC)	UNIV CALIFORNIA.				
XX	Scott MP,	Goodrich LV, Johnson RL, Epstein E;				
XX	WPI;	2004-041193/04.				
DR	N-PSDB;	AD548988.				
XX	Phenotyping the patched status of a cell for diagnosing a genetic					
PT	predisposition for a tumor comprises detecting the presence or absence of					
PT	a genetic lesion having aberrant modification, mutation or mis-expression					
PT	of the patched gene.					
XX	Disclosure; SEQ ID NO 19; 60pp; English.					
XX	The invention describes an assay for phenotyping the patched status of a					
CC	cell comprising detecting in a sample of mammalian cells the presence or					
CC	absence of a genetic lesion having aberrant modification or mutation of a					
CC	patched gene or mis-expression of the patched gene. The assay is useful					
CC	for diagnosing a genetic predisposition of an animal, e.g. basal cell					
CC	nevus syndrome, predisposition for developing tumour, i.e. carcinoma.					
CC	meningioma, medulloma, or fibroma. A genetic construct encoding a patched					
CC	polypeptide is used to treat an animal having a disorder comprising loss					
CC	of function of a wild-type patched gene, such as cancer, and can enhance					
CC	patch function in e.g. wound healing and aging. This sequence is encoded					
CC	by the human patched gene.					
XX	SQ Sequence 1447 AA;					
Query Match 9.4%; Score 651.5; DB 8; Length 1447;						
Best Local Similarity 22.0%; Pred. No. 2.4e-50;						
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;						
QY	225	PGQAVSGIOPFNEGVARCNESQGDVATCSQDCAASCAPAIARPOLDSTFYLGQMPGS	284			
Db	24	PGRPAGGRRRTGRLRAAADPRD-----YLHRPSYCDAAFALEQI----	65			
QY	285	LVLIILICSPVAVNTILLVGRVAPARDKSMWDPK-----KGTSLSDKLSFSTHTLL	337			
Db	66	-----SKGATGRKAPLMRAKFORLLFKLGCYIQNC	98			
QY	338	GQFFQGWGTWASWPLATILVLSVPVVALAAGLVFETLTDVPELWSAPNSQASEKAFH	397			
Db	99	GKF-----LVVGLLI FGAFVGLKAAANLETNVEELWVEVGVRVRELNVT	143			
QY	398	DQHFGPPFRTN-QVILFAPNRSRYDSLLGPKNFSGILDLLELLE--LQERLRLH	454			
Db	144	RQKIGEEAMFNQPMIQTPEEG-----ANVLTEALLQHLDSALQSRVHV	190			
QY	455	QWVSEAOENISLQDICVAP-----LNPNTSLYDCCINSLLQYFQNNRTLLTLTA	505			
Db	191	YMYN-----RQWLEHLYKSGELITETGYMQIIEYLYPCLITPDLDFWEGAKLSGTA	246			
QY	506	NOTLMGQ-----TSQVDWKDHFY-----CAN-----	527			
Db	247	--YLLGKPLRWTFDPLEFLEELKKINVQVDSWEEMLNKAHVGHGYMDRPCINPADPC	304			
QY	528	-----APL-----TFKDG-----ALALSCMA	544			
Db	305	PATAPNKNSTKPLDMALVNLGCHLSRKRYMHQBELVGGTVKNTKGLVSAHALQTMF	364			
QY	545	DYGAPVFPPLAIGGVKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFO	604			
Db	365	QLMTPQMY---EHFAGYFY-----VSHINWNE---DKAAAILLEAWQRTYVEVHVQS	411			
QY	605	RRMAGMFQVTFTAERSLEDEINRTAEDLPFATSPVIVIFLISALGSSYSSNRVWVDS	664			
Db	412	AQNSTQKVLSTFT--TTLDDILKSFSDSVIRVASGYLLMAYACLTM---LRWD---CSKS	465			
QY	665	KATLGLGGVAVVLGVAWMAAGFFSYLGRSSVLVQVVPFVLSVGADNIFIVLEVYQRL	724			

Db 366 DY-----VSHINWNE-----DRAAAILEAWQRTVEVHQSVAPNPTQKVLPTFT-TTLD 414
Qy 623 DEINRTTAEDLPPIFATSVIVIFLXISALGSSWSRVMVDSKATLGLGGVAVVLGAVMA 682
Db 415 DILKSFSDSVIRVASGYLLMLAVACLTW-----LRWD--CSKSQAGAVLAGVLLVALSVAA 469
Qy 683 AMGFYSYLGRSSIVILQVQVFLVLSGADNIFIVLEYQRLPRRPGEPREHVHIGRALGR 742
Db 470 GLGLCSLIGISFNAATQVLPFLALGVDDVDFLLAHAFSETGQNKRIPEPDRGTGECLKR 529
Qy 743 VAPSMMLCSLSEATCFFGALGATPMBAVNTFALTGSLAVILDFLOMSAFVALLSLDSKQ 802
Db 530 TGASVALTSISNVTAFMAALIPALRAFSLQAAVVVFNFMVLLFIFFAILSMIDLYRR 589
Qy 803 EASRLDVCCC-----VKQEL-----PPGQEGEL----- 827
Db 590 EDRRLDIFCCFTSPCVSRVIQVEPOAYTEPHSNTRYSPPPYTSHPAFETHITMQSTVQ 649
Qy 828 ----- 827
Db 650 LREYDPHTHYVYTTAERSEISVQPVTVTQDNLSCQSPSTSTRDLSQFSDSSLHCL 709
Qy 828 -----LLGPFQKAYAPFLHWTIRGVLLFLALFGVSLYSCHISVGLDQELALPK 879
Db 710 EPPCTKWTLSIPAERHYAPFLKPKAKVVVILLFGLLGVSLYGTTRVRDGLDLTDIVPR 769
Qy 880 DSYLDYDFLFLNRYFEVGPVYVTTLVGNFSSSEAGMNAICSSAGCANNFSFTQKIQYAT- 938
Db 770 ETRYDFIAQPKYFSF-----YNM-----YIVTKADYENI 801
Qy 939 -----EPPEQSYLAIPASS-----WVDDFDLW-----TPSS--- 965
Db 802 QHLLYDLHKSFSNVKVMLEENKQPLQMWLHYFRDMLQGLQDAFSDWETGRIMPNYKN 861
Qy 966 -----CRLYISGNKDKFCPSTVNSLNCIKNCSITMGVSRSVQFHKYL-PWFLN 1017
Db 862 GSDDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWVSN 915
Qy 1018 D-----RPNIKCPKGGAAAY--STSVNLTSDGQVLASRFMAYHKLNQSD 1061
Db 916 DPVAYAAQANIRPHRPEWHDK--ADYWPETRLRIPAAEPIEYAAQFPFLYNGLRDTS 972
Qy 1062 YTEALSAARELANITA-DLRKVPGTDPAPFVPPYITITNVFEQYLTITLPEGLFMLSCL 1120
Db 973 FVEAIEKVRVICNNYTSGLSSYPNG-----YPF-----LFWEQYISLRHWLLLSISVVL 1022
Qy 1121 VPTFAVSCLLGLDLRSGLNLLSIYVILVDTVGFNALWDISYNAVSLINLVSAGVMSVE 1180
Db 1023 ACTFLVCAVFLNPNWTAGII-VNVLALMTVELFGMMGLIGIKLSAVPVVILLIASVGIGVE 1081
Qy 1181 FVSHITRSF-AISTKPTWLERAKEATISNGSAVFAGVAMTNLPGLVLVLGAKAQIQT 1238
Db 1082 FTVHVALAFITAGDKN--HRAWLALEHMFAPVLDG-AVSTLLGVLMLAGSEDFIVRY 1137
Qy 1239 FFRNLNLTLLGLLHGLVFLPVLTSVG--PDVNPALAEQ-----KRAEBAVAVMWASCP 1293
Db 1138 FPAVLAITLVGLNGLVLLPVLSPFGPCPEVSPANGNLRLPTSPPEPPPVVREAVPP 1197
Qy 1294 NHRSVS-TADNIYVNHSPGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISELRYEAOQAG 1234

RESULT 44
AAW72968
ID AAW72968 standard; protein; 1434 AA.
XX AC
XX AC
XX AC
DT 26-JAN-1999 (first entry)
XX DE
DE Mouse patched protein.

XX Mouse; patched gene; diagnosis; treatment; developmental disorder;
KW cancer; healing; injured tissue; ptc; spina bifida; Wnt-1 oncogene;
KW sperm production; gene therapy.
OS Mus sp.
XX US5837538-A.
XX 17-NOV-1998.
XX 06-OCT-1995; 95US-00540406.
XX 07-OCT-1994; 94US-00319745.
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX Scott MP, Goodrich LV, Johnson RL;
PI WPI; 1999-023461/02.
DR N-PSDB; AAV64092.
XX Nucleic acid encoding vertebrate patched protein and related
transformants - used to express poly.peptide(s), useful for diagnosis and
treatment of developmental disorders or cancer, and in healing of injured
tissue.
XX Claim 23; Col 49-56; 38pp; English.
XX The present sequence represents the mouse patched (ptc) protein. Cells
containing and expressing the ptc gene are used for the recombinant
production of the protein. These in turn are useful: (i) for generating
antibodies (Ab); and (ii) to screen for specific-binding ligands
(potential therapeutic agonists and antagonists). The ptc gene, or its
fragments, are used to isolate related sequences from other mammals; to
identify mutations (particularly those associated with genetic diseases
such as spina bifida and other developmental disorders); to monitor
expression levels in testis (to determine relationship with sperm
production) and to isolate 5'-non-coding sequences (used to study
embryonic development and to provide regulated expression of proteins).
The complete gene can be used in gene therapy, including expression of
antisense molecules, and to generate transgenic animals for studies of
embryonic development. Ab are used diagnostically to determine the ptc
protein on cell surfaces and as competitive inhibitors of signal
transduction through the ptc ligand. Cells that have been engineered to
express the ptc protein can be used to promote regrowth and healing of
damaged tissue (e.g. growth of new teeth) and regulation of the ptc
protein expression may be useful in cancer treatment (it may control the
Wnt-1 oncogene)
XX Sequence 1434 AA;
SQ
Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 7e-50;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
Qy 356 LVLSVIVPVALAAGLVFTTELTDVPELWSAPNSQARSEKAFHDQFHPFRFN-QVILFA 414
Db 88 LVVGLLIIFGAFVAGLKAANLETNVELWVEGVRVSRLENTYTRQIGEAEMFNPLMIQT 147
Qy 415 PNRSSRYRVDLLGPKNFSGILDLLLELLE--LQERLRLHQVMSPEAQRNLSLDICY 472
Db 148 PKEEG-----ANVLTEALLQHLDSALQASRVHYMYN-----RQWKLEHLCY 190
Qy 473 AP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLLTANQTLMOG----- 512
Db 191 KSGELITETGYMDQIIIEYLYPLCIITPLDCFWEGAKLQSGTA--YLLGPKPLRWTFDPL 248
Qy 513 -----TSQVDWKDHELY-----CAN-----APL----- 530
Db 249 EFLEELKKINYQVDSWEEMLNKAEGVGHGYMDRPNLPADPCDPATAPNKNSTKPLDVALV 308
Qy 531 -----TFKQGT-----ALALSCMADYAGAPVFPFLAIGYKKG 562

Db 309 LNGCGGLGRKYMHWQEEELIVGTGVRNATGKLSAHALQTMFQMTPKQMYEHFRGY--- 365
QY 563 DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFQVTFPTAERSLE 622
Db 366 DY-----VSHINWNE-----DRAAAILEAQRYVEVHVQSVAPNSTQKVLPTT-TTLD 414
QY 623 DEINRTAEDLPFPATSYIVIFLYISLALGSYSSSRVMVDSKATGLGGVAVVLGAVMA 682
Db 415 DILKSPFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKQCAVGLAGVLLVALSVAA 469
QY 683 AMGPFYVIGIRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGEPREVHIGRALGR 742
Db 470 GLGLCSLIGISENAATQVLFALGVGDVDFELLAFAFSETGQNKRIIPFEDRTGCLXR 529
QY 743 VAPSMLLCSLSEALCFGLAGTTPMPAVRTFALTGLAVILDPLLOMSAFVALLSLDSKRO 802
Db 530 TGAASVLTISNVTAFMAALPIPALRAFSLQAIVVVVFNFAMVLLIPFALLSMDLYRR 589
QY 803 EASRLDVCCC-----VKQDEL-----PPGQGEGL----- 827
Db 590 EDRRLDIFCFTSPCVSRVIOEQAYTEPHSNTRYSPPPVYTSFSAHETHITMQSTVQ 649
QY 828 ----- 827
Db 650 LRTEYDPHTHYVYTTAEPREISVQPVTVTDNLSCQSPSTSTRDLSQFSDSLHCL 709
QY 828 -----LLGFQKAYAPFLHWTIRGVVLLFLALFGVSLYSMSCHISVGLDQELALPK 879
Db 710 EPPCTKWTLSFAEKYAPFLKPKAKVVILLFLGLGVSLYGTTRVRDGLDLDIVPR 769
QY 880 DSVLLDYFLNRYFEVGAFFVFTVTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYAT- 938
Db 770 ETRVEDFIAAQKYPF-----YNN-----YIVTQKADYPI 801
QY 939 -----EPEQSYLAIPAS-----WDFIDWL-----TPSS----- 965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQMWLHYFRDMLQLODAFSDMETGRIMPNNYKN 861
QY 966 -----CCRLYISGNPKDKPCSTVNSLCKNCSITWGSVRPSVEQPHKYL-PWFLN 1017
Db 862 GSDGVLAYKLIVQTSRDK--PIDISQUT--KQRLVDADGIINPSA--FYIYLTAWVSN 915
QY 1018 D-----RPNIKCPKGGLAAY--STSVNLTSQGVQVLSRFRMAYHKLNKNSOD 1061
Db 916 DPVAYAASQANIRPHPEWVHDK---ADYMPETRLRIPAAEPIEYQAQFPFYINGLEDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPQTDPAFEVFPYTTITNVFYEQYLTILPEGLFMLSCL 1120
Db 973 FVEAIEKVRVICNNYTSLSGSSYPNG-----YFP-----LFWEQYISLRHMLLSVVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLLNLLSIVMLIVDTVGFMAWDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCVFLNPTWAGII-VWVLAJMTVELFGMGLIGIKLSAVPVVILLASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTMLERAKATISMSGSAFAGVAMTNLPGLIVLGLAKAOLQIF 1238
Db 1082 FTVHVALAFLTAIGDKN---HRAMLALHEMFAPVLGD-AVSTLLGVLMLAGSEFPFIVRY 1137
QY 1239 PFRMLLITLLGLHGLVPLVLSVVG--PDVNPALAEQ---KRAEAAVAVMVASCP 1293
Db 1138 PFAVLAAILTVGLVNLGLVLLPFLSFGPCPEVSPANGLNRLPTSPPEPPSVVRFAVFP 1197
QY 1294 NHPSRVS-TADNINYHNSPEGSI-----KKG 1319
Db 1198 GHTNNGSDSDSEYSSQITVSGISBELROYEAQQGAG 1234

RESULT 45

AAB67159

ID AAB67159 standard; protein; 1434 AA.

XX

AC AAB67159;

XX 12-APR-2001 (first entry)
DT Murine patched protein.
DE Mouse; patched; PTC; segment polarity; limb patterning; development;
KW hedgehog; antibody; human.
XX Mus sp.
XX US6172200-B1.
XX 09-JAN-2001.
XX 20-OCT-1997; 97US-00954668.
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX (STRD) UNIV LELAND S STANFORD.
XX Scott MP, Goodrich LV, Johnson RL;
XX WPI; 2001-136884/14.
XX N-PSDB; AAF32180.
XX Novel monoclonal antibody useful in diagnostic assays for detection of
PT presence of protein on surface of cells specifically binds to naturally
PT occurring patched protein, other than Drosophila patched protein.
XX Disclosure; Col 47-54; 39pp; English.
XX The present invention provides a monoclonal antibody which specifically
CC binds to a patched protein (PTC) other than that from Drosophila. Also
CC given are the protein and coding sequences of patched from the beetle,
CC mouse, butterfly, fruit fly, mosquito and human. Patched is a segment
CC polarity gene involved in limb patterning. The sequences can be used to
CC study development and to isolate the patched ligand, hedgehog. In
CC addition, antibodies can be used to detect the PTC protein on cell
CC surfaces or to inhibit the transduction of signal by the PTC ligand by
CC competing for its binding site
XX Sequence 1434 AA;
Query Match 9.4%; Score 646.5; DB 4; Length 1434;
Best Local Similarity 22.9%; Pred. No. 78-50;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
QY 356 LVLSVIPVVALAAGLVFTTELTTDPVELMSAPNSQARSEKAFHDQHFPPFRTN-QVILTA 414
Db 88 LVVGLLIFGAPAVGLKAANLETNVELWVEVGRVSRRELNTYTRQKIGBEAMFNPQLMIQT 147
QY 415 PNRSSRYVDSLLGPNKPSGILDLDLLELE--LQERLRHLQVMSPEAQRNLSIDICY 472
Db 148 PKEEG-----ANVLTTEALLQHLDSALQASRVHYVMYN---RQWKLEHLCY 190
QY 473 AP-----LNPNTSLYDCINSILLOYFQNNRTLLLLTANOTLMGQ----- 512
Db 191 KSGELITETGMDQIIEYLYPLCLITPLDCFWEGAKLOSHTA--YLLGKPLPLRTNFDPL 248
QY 513 -----TSQVMDKHFLY-----CAN-----APL----- 530
Db 249 EFLBELKXINQVDSWEEMLNKAEVGHGYMDRPNCLPADPCPATPNKNSKTKPLDVALV 308
QY 531 -----TPKQGT-----ALALSCWADYGAIPVFPFLAIGYKKG 562
Db 309 LNGCGGLSRKYMHWQEEELIVGTGVRNATGKLSAHALQTMFQMTPKQMYEHFRGY--- 365
QY 563 DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFQVTFPTAERSLE 622
Db 366 DY-----VSHINWNE-----DRAAAILEAQRYVEVHVQSVAPNSTQKVLPTT-TTLD 414
QY 623 DEINRTAEDLPFPATSYIVIFLYISLALGSYSSSRVMVDSKATGLGGVAVVLGAVMA 682

[illegible]

```
QY 563 DYSEBALIWTSLSNYPAGDPRPLAQAKLWERAFLEEMRAFORRMAGMFOVFTTFAERSLE 622
DB 366 DY-----YSHINWNE-----DRAAAILEAQRTYVEVHQSVAPNSTQKVLFPFT--TTLD 414
QY 623 DEINRTTASDLPIFATSYIVIFLYISLALGSYSRVMVDKATGLGGVAVLGAWNA 682
DB 415 DILKSPSDSVIRVASGYLLMLAYACLTW---LRWD--CSKQAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVLOVVPVLVSGADNIFIFVLEYQRLPRRPGEPREHIGRALGR 742
DB 470 GLGLCSLIGISFNAATTVQLPFLALGVDDVFLLAHAFSETGQNKRIFFEDRTGECLGR 529
QY 743 VAPSMMLCSLSRAICFFICALTPMPAVRTFALTSLGLAVLDFLLQMSAFVALLSDSKRQ 802
DB 530 TGSVALTISINVTAFPAALPIPALRAFSLQAQAVVVVFNAFVILLIIPAILSMDLVYR 589
QY 803 EASRLDVCCC-----VKQEL-----PPQGEGL-----827
DB 590 EDRLDIFCCFTSPCVSRVIOEQAYTEPHNTRYSPPPPTSHSFAHETHITWQSTVQ 649
QY 828 -----827
DB 650 LRTEYDPHTHYVYTTAEPSEISVQPVTVTQDNLSCQSPSTSSTRDLLSQSDSLHCL 709
QY 828 -----LLGFQKAYAPFLHWTITRGVLLLFALFGVLSYSMCHISVGLDQELALPK 879
DB 710 BEPCTKWTLSSPAETHYAPFLKPKAKVVVILLFLGLLGVSLYGTTRVADGLDLTDIVPR 769
QY 880 DSYLLDYFLFNRYFEVGPVTVTTLGYNFSEAGMNAICSSAGCNPFSTQKIQYAT- 938
DB 770 ETREYDFIAQPKYFSF-----YNM-----YIVTQKADYPMI 801
QY 939 -----EPFEQSYLAIPASS-----WDDFDIDL-----TPSS---965
DB 802 QHLLYDLHKSFNKVKYMLEENKQIPOMWLYFRDLWLOGLQDAPSDWETGRIMPNYKN 861
QY 966 -----CRLVISGNKDKFCPTVNSLNCMKMSITMGSVRPSVQPHKYL--PWFLN 1017
DB 862 GSDDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWSN 915
QY 1018 D-----RPNKCPKGLAAY--STSVNLSDGQVLASRPMAYHKPLKNSQD 1061
DB 916 DPVAAASQANTRPHRPVWHDK--ADYMPETRLRIIPAAEPYEAQFPFYINGLRLDTS 972
QY 1062 YTEALRAARELANITA--DLRKVPQGTDAPEFVFPYITITNVFEQYLTILPEGLFMLSCL 1120
DB 973 FVEALEKVRVICNWTSLGLSYFNG-----YFP-----LFEQYISLHMLLSISVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFPMALWDISYNVSLNLSVAYGMSVE 1180
DB 1023 ACTFLVCVAVFLNPTAGII--VMVLALMTVELFGMGLIGIKLSAVPVVILLIASVIGVE 1081
QY 1181 FVSHITRSF--AISTKPTWLERAKAETISGSAVAGVAMTNLPGILVLGLAKAQLIQIF 1238
DB 1082 FTVHVALLAFLTAIGDKN---HRAMLALEHMFAPVLDG--AVSTLLGLVLMAGSEDFIVRY 1137
QY 1239 PPRMLLITLLGLLHGLFLPVLISVG--PDVNPALALEQ---KRAEVAVAAMVASCP 1293
DB 1138 FPAVALITVLGLVNLGLVLLPILLSFFGCPCEVSPANGLNRLPTSPBPSPSVRFVAPP 1197
QY 1294 NHPSRVS--TADNIYNHSPGSI-----KGAG 1319
DB 1198 GHTNNGSDSDSEYSSQITVSGISELRQYEAQOGAG 1234
```

RESULT 47

ABU62271

ID ABU62271 standard; protein; 1434 AA.

XX AC

ABU62271;

XX 25-AUG-2003 (first entry)

XX

DE Mouse patched gene PTC product.

XX Mouse; patched gene; PTC; hedgehog protein; gene therapy.

XX Mus sp.

XX US2003032085-A1.

XX 13-FEB-2003.

XX 20-OCT-1997; 97US-00954701.

XX 07-OCT-1994; 94US-00319745.

XX 06-OCT-1995; 95US-00540406.

XX (SCOT/) SCOTT M P.

XX (GOOD/) GOODRICH L V.

XX (JOHN/) JOHNSON R L.

XX Scott MP, Goodrich LV, Johnson RL;

XX WPI: 2003-492065/46.

XX N-PSDB; ACA62834.

XX New DNA sequence other than present in a chromosome encoding patched gene
other than Drosophila patched gene, useful for producing transgenic
laboratory animals and to knock out patched protein in embryonic stem
cells.

XX Disclosure; Page 8-10; 40pp; English.

CC The invention relates to a DNA sequence other than present in chromosome
encoding a patched (ptc) gene other than Drosophila patched gene. A cell
expressing the DNA sequence is useful for producing patched protein, by
growing the cell expressing the DNA sequence, where the patched protein
is expressed and isolating the patched protein free of other proteins.

CC The cell expressing the DNA sequence is also useful for screening

CC candidate compounds for binding affinity to the patched protein, by

CC combining the candidate protein with the cell expressing the DNA sequence

CC where the DNA sequence comprises the entire coding sequence under the

CC transcriptional regulation of the transcriptional initiation region and a

CC patched protein in the cell and assaying for the binding of the candidate

CC compound to the patched protein. The above cell is useful for screening

CC candidate compounds for agonist activity with the patched protein. The

CC DNA sequence is useful for producing all or portions of the patched

CC protein, as probes for research, diagnosis, binding of hedgehog protein

CC for its isolation and purification and in gene therapy. The DNA sequence

CC is also useful as primers for investigating other species and for

CC isolating genes from various mammalian sources of interest, particularly

CC from humans or from domestic animals. The DNA sequence is further useful

CC for preparing transgenic laboratory animals and to knock out the PTC

CC protein in the embryonic stem cells, so as to produce hosts with single

CC functional patched gene. The present sequence represents the amino acid

CC sequence of the mouse patched gene PTC product

XX

SQ Sequence 1434 AA;

Query Match 9.4%; Score 646.5; DB 7; Length 1434;

Best Local Similarity 22.9%; Pred. No. 7e-50;

Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

QY 356 LVLSVIPVVALAAGLVFTTELTDPVELWSAPNSQARSEKAFHQHGFPRTN-QVILTA 414

DB 88 LVWGLLTGPAFVGLKAALETNVEELVVRVGGVRSRELTNRQKIGSEAMFNQMIQT 147

QY 415 PNRSSRYDLSLLGPKNPSGILDLDLLELE--LQERLRHQVMSPEAQNISLDICY 472

DB 148 PKREG-----ANVLTTEALLQHLDSALQASRVHVMYN-----RQWKLEHLCY 190

QY 473 AP-----INPNTSLYDCCINSLLQYFQNNRTLLLLTANQTLMOQ-----512

DB 191 KSGELITGTGYMDQIIIEYLPCLITPLDCEWEGKLGSGTA--YLLGKPLPLWTNPDPL 248

QY 513 -----TSQVDWKDHFY-----CAN-----APL----- 530
Db 249 EFLEELKINYQVDSWEMLNKAEGVGHGYMDRCLNADPADPCDPATAPNKNSTKPLDVALV 308
QY 531 -----TFKDG-----ALALSCADYDAPVFFFLAIGYKKG 562
Db 309 LGGCQGLSRKMYHQBELIVGGTVKNATKLSAHALQTMFLMTPKMYEHFRGY--- 365
QY 563 DYSABALIMTFSLNYPAGDPLRAQAKLMEEAFLBEMRAFORRMAGMFOVTFIATERSLE 622
Db 366 DY-----VSHINNE-----DRAAAILLEAQRTVEVHVQSVAPNSTOKVLPFTT-TTLD 414
QY 623 DEINRTAEDLPIFATSYIVIFYLSIALGSSYSWSRVMVDSKATLGGVAVVLGAVMA 682
Db 415 DILKSFSVSVIRVASGYLLMLAYACLTM---LRWD--CSKSQGAVALGLVLLVALSVA 469
QY 683 AMGFPSYLGIRSSLVILQVDFVLVSUGADNIFIFVLEYQBLPRPGEPREVHIGRALGR 742
Db 470 GLGLCSLIGISFNAAATQVLPFLALGVGDVDFLLAHAFSETGQNKRIFFEDRTGECLKR 529
QY 743 VAPSMMLCSLSEACIFELGALTEPMPAVRTFALTGSLAVILDFLLQMSAFVALLSLDSKRQ 802
Db 530 TGASVALTSISNVTAFFMAALIPILALRAFSLQRAVVVFNFAVLLIFPAILSMWLYRR 589
QY 803 EASRLDVCCC-----VKQEL-----PPGQEGEL----- 827
Db 590 EDRLDIFCCFTSPCVSRVIOVEPQAYTEPHSNTRYSPPPVYTSHPAHETHITMQSTVQ 649
QY 828 ----- 827
Db 650 LRTEYDPHTHYTTAEPRSEISQVPTVTQNLSCQSPSTSTRDLLSFSSSLHCL 709
QY 828 -----LLGFFOKAYAPFLHWHITRGVLLFLALPGVLSYMSCHISVGLDQBELALPK 879
Db 710 EPPCTKWLTSSFAEKYAPFLKPKAKVIVLLFGLLGVSLYGTTRVDRDGLDLTDIVER 769
QY 880 DSYLDYFLNRYFEVAGPVVFTVLGVNFSSEAGMNAICSSACGNPFSFTKQIYAT- 938
Db 770 ETRYDFIAAQPKYFSF-----YNM-----YIVTKADYPNI 801
QY 939 -----EPFQSYLAIPASS-----WDDPFIWL-----TPSS--- 965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQMWLHFRDMLQGLQDAPSDMETGRIMNNYKN 861
QY 966 -----CCRLYISGPNKDKFCPSTVNSLCLNKNCSITMGSVRPSVEQFHKL-PWFLN 1017
Db 862 GSDGVLAYKLLVOTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWVSN 915
QY 1018 D-----RPNTKCPKGGLAAY--STSVNLTSDGQVVLASRFMAYHKPLKNSOD 1061
Db 916 DPVAYAASQANIRPHRPEWVHDK---ADYMPETRLRIPAAEPIEYAOFPFYLINGURDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPGTDPAPVEVPYITITNVFYBOYLITLPEGLFMLSCL 1120
Db 973 FVEAIEKRVICNNYTSGLSYPNG-----FPF-----LFWQYISLHWHLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLRGLNLLSIVMLIVDTVGFMAWMDISYNAVSLINLSAVGMSVE 1180
Db 1023 ACTFLVCVAVFLNPTWAGII-VWVYALMTVELFGMGLIGKLSAVPVVILIASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIIVLGLAKAQLIQIF 1238
Db 1082 FTVHVALAFLTAIGDKN---HRAMLALEHMFAPVLDG-AVSTLLGLVLMLAGSDFDIVRY 1137
QY 1239 FFRNLALITLLGLHGLVFLPVILSVYG--PDVNPALALRQ---KRAEAEVAAVAMVASCP 1293
Db 1138 FFVAILITLVGLVNLGLVLLPVLSFPGCPCEVSPANGLANRLPTPSPPPSVVRFAVPP 1197
QY 1294 NHPSRVS--TADNIYVNHHSFEGSI-----KGAG 1319
Db 1198 GHTNNGSDSSDSEYSSQTTVSGISEELRQYEAQQAG 1234

RESULT 48
ADE94215
ID ADE94215 standard; protein; 1434 AA.
XX
AC ADE94215;
XX
DT 12-FEB-2004 (first entry)
XX
XX Murine ptc polypeptide.
XX Mouse; patched gene; ptc; developmental abnormality;
XX proliferative disorder; tumour; oncogenic patched mutation;
XX patched-associated phenotype; basal cell nevus syndrome; BCNS; carcinoma;
XX meningioma; fibroma; cancer.
XX Mus sp.
XX US6551782-B1.
XX
XX 22-APR-2003.
XX
XX 28-NOV-2000; 2000US-00724631.
XX
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX 31-MAY-1996; 96US-00656055.
XX 22-AUG-1997; 97US-00918658.
XX
XX (STRD) UNIV STANFORD.
XX (REGC) UNIV CALIFORNIA.
XX
XX Scott MP, Goodrich LV, Johnson RL, Epstein E;
XX
XX WPI; 2003-615308/58.
XX N-ESDB; ADE94214.
XX
XX Assay for phenotyping patched status of cell in mammalian (preferably
XX human) cell sample, involves detecting presence or absence of aberrant
XX modification or mutation of a patched gene, and mis-expression of patched
XX gene.
XX
XX Disclosure; SEQ ID NO 10; 57pp; English.
XX
XX The invention relates to an assay for phenotyping the patched status of a
XX cell, involving detecting in vitro, in a sample of mammalian cells, the
XX presence or absence of a genetic lesion in a patched gene (ptc)
XX characterised by at least one of an aberrant modification or mutation of
XX a patched gene or mis-expression of the patched gene. The invention also
XX relates to a method for diagnosing a genetic predisposition of an animal
XX to a developmental abnormality or a proliferative disorder, where the
XX developmental abnormality or proliferative disorder correlates with
XX aberrant expression or activity of a patched gene or gene product,
XX involving detecting in vitro the presence of a predisposing mutation in a
XX patched gene in cells of the animal. Characterising the phenotype of a
XX tumour, involves detecting the presence of an oncogenic patched mutation
XX in cells of the tumour, where the presence of the oncogenic mutation
XX indicates that the tumour has a patched-associated phenotype. The assay
XX is useful for phenotyping the patched status of the cell in a mammalian
XX cell sample obtained from a human patient. The disorders that can be
XX detected include basal cell nevus syndrome (BCNS), carcinomas,
XX meningiomas and fibromas. This sequence represents a murine ptc
XX polypeptide of the invention.
XX
XX Sequence 1434 AA;
XX
XX Query Match 9.4%; Score 646.5; DB 7; Length 1434;
XX Best Local Similarity 22.9%; Pred. No. 78-50; Indels 363; Gaps 44;
XX Matches 283; Conservative 171; Mismatches 420;
QY 356 LVLSVPPVVALAAGLVFTLTTPDVELWSAPNSQARSEKAFHDQHFPPFRN-QVILTA 414
Db 88 LVVGLLIIFGAFVGLKAALEINVELWVEGVSRVSELNTRYTRQKIGEBAMFNPQLMIQT 147

Qy	415	PNRSRYDSSLGKPNFSGILDJDLLELE--LQERLRHLQWMSPEAQRNISLODICY	472
Db	148	PKBEG--ANVLTEALLQHLDSALQASRVHVYMN--ROWKLEHLCY	190
Qy	473	AP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTANCTLMGQ-----	512
Db	191	KGELITETGYMDQII EYLPCLLIITPLDCFWEGAKLQSGTA--YLLGKPLRLWNPDPDL	248
Qy	513	-----TSQVDWKDHFY-----CAN-----APL-----	530
Db	249	EFLEELKKINYQVDSWEEMLNKA EVGHYGDMR PCLNPADPCDPATAPKNKSTRPLDVALV	308
Qy	531	-----TFKQGT-----ALALSCWADYGAPVPFPFLAIGGVKKG	562
Db	309	LNGCQGLSRKYMWOBELIVGGTVKNAKTGLVSAHALQTMFQMLTPKQMYEHRFGV---	365
Qy	563	DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLBEWRAPQORMAGMPOVTTABESLE	622
Db	366	DY-----VSHINNE-----DRAAAII EANQRTTVEVHVOSVAPNSQKVLPFTT--TILD	414
Qy	623	DEINRTAEDLPIFATSYIVIFLYISLALGSYSWSRVMVDSKATLGLGGVAVVLGAVMA	682
Db	415	DILKSFSDSVIRVASGYLLMLAVACLTM--LRWD--CSKSGAGVGLAGVLLVALSVAA	469
Qy	683	AMGFPSYLGIRSSLIIVQVPFLVLSVGADNIFIFVLEYQRLPRRPOPEPREVHIGRALGR	742
Db	470	GLGLCSLIGISFNAAATTQVLPFLGALGVGVDDVFLLAHAFSETGQNKRIPFEDRGTCECLR	529
Qy	743	VAPSMLLCSLSEALCFPLGALTTPMAVRTALTLSGLAVILDFLQLMSAFVALLSLDSKQ	802
Db	530	TGASVALTSSNTVAFMAALIPALRAVSLQAAVVVFVFAVLLIFPAILSMDLYRR	589
Qy	803	EASRLDVCCC-----VKPQEL-----PPPGQEGEL-----	827
Db	590	EDRLDIFCCFTSPCVSRKVIQVEPQAYTEPHSNTRYSPPPPYTSHSPAETHITMQSTVQ	649
Qy	828	-----TPSS-----	965
Db	650	LRTEYDPTHVYTTABRSEISVQPVTVTDNLSCQSPSTSTRDLLSQFSDSSLHLCL	709
Qy	828	-----LLGFFOKAYAPFLHWHITRGVLLLFLALFGVSLYSMSCHISVGLGQELALPK	879
Db	710	EPPCTKWTLSSFAEKHYAPFLPKKAKVWILLFLGLLGSVLYGTTVRGGLDLTDIVPR	769
Qy	880	DSYLLDYFLNRYRFEVGAPVYFVTTLGYNFPSSBAGMNAICSSAGCNFNFTQKIQYAT-	938
Db	770	ETREYDFTAAQFKYFSF-----YNN-----YIVTQKADYPNI	801
Qy	939	-----EPPEQSYLAIPASS-----WVDDDFDML-----TPSS---	965
Db	802	QHLLYDLHKFSNRYKMYLNEENKOLPQWHLHYFRDLWLOGLQADPDSWETGRIMPNNYKN	861
Qy	966	-----CCRLIYISGPNKDFPCPSTVNSLNCNKMCSITMGSRVPSVEQFHXYL--PWFLN	1017
Db	862	GSDGVLAYKLLVGTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FVILYTAWSVN	915
Qy	1018	D-----RPNIKCPKGGLAAY--STSVNLTSGQVLASRFMAYHPKLKNSQD	1061
Db	916	DPVAYAAQANIRHRPEWHDK---ADYMPETRLRIPAAEPIEYAFQFPYFLNGLRDTSD	972
Qy	1062	YTEALRAARELAANITA-DLRKVPGTDPAPRVPRTITNRYEORYLTILPEGLFMLSCLCL	1120
Db	973	FVEAIEKRVICNNYTSILGLSSYPNG-----YFF-----LFWEQYISERHWLLLSISVVL	1022
Qy	1121	VPTFAVSCLLGLDLRSLGLMLLSIVMLIVDVTGFMALMDWISYNAVSLINLVSAGVMSVE	1180
Db	1023	ACTFLVCVAVFLNPNWTAGI--VMVLA MTLVELFGMWGLIGKLSAVPVVILLIASVGIGVE	1081
Qy	1181	FVSHITRSP--AISTKPTWLERAKENTISMGSAVPAGVAMTNLPGLIUVGLAKAQILQIF	1238
Db	1082	FTHVVALAFTPAIGDKN---HRAMLAEHMFAPVLDG-AVSTLLGVLMLAGSFPDFTVRY	1137

Qy	1239	PFRNLNLTLLGLLHGLVFLPVLISVVG--PDVNPALALRQ---KRAEEAAVAAMVASCP	1299
Db	1138	FFAVLAAILTVLGLVGLVLLPVLISFPGCPCEVSPANGLNRLPTSPSPPPSVVRFAVPP	1197
Qy	1294	NHPSRVV-TADNLYVNHSPFGSI-----KGAG	1319
Db	1198	GHTNNGSDSDSEYSSQTTVTSGISELRQYEAQOGAG	1234
RESULT	49		
ADH62722			
ID	ADH62722	standard; protein; 1434 AA.	
XX	ADH62722;		
XX	AC		
XX	AC		
DT	25-MAR-2004	(first entry)	
XX	Mouse patched	(mptc) protein.	
DE			
XX	PTC; patched;	embryonic development; cellular regulation;	
KW		signal transduction; mouse.	
KW			
XX			
OS	Mus sp.		
XX	US2003148388-A1.		
FN			
XX	07-AUG-2003.		
PD			
XX	03-JAN-2001;	2001US-00754032.	
PF			
XX	07-OCT-1994;	94US-00319745.	
PR	06-OCT-1995;	95US-00540406.	
PR	20-OCT-1997;	97US-00954668.	
XX	(STRD)	UNIV STANFORD.	
PA			
XX	Scott MP,	Goodrich LV, Johnson RL;	
PI			
XX	WPI; 2003-897566/82.		
DR	N-PSDB; ADH62721.		
DR			
XX	New DNA sequences encoding patched protein from organisms including human		
PT	and mouse are used to elucidate embryonic development and cellular		
PT	regulation associated with patched signal transduction and identify		
PT	patched ligands.		
XX	Disclosure; SEQ ID NO 10; 40pp; English.		
FS			
XX	The present invention relates to a DNA sequence encoding a patched (PTC)		
CC	gene. The invention is used to elucidate embryonic development, cellular		
CC	regulation associated with signal transduction by the patched gene,		
CC	identify agonists and antagonists of that signal transduction, identify		
CC	and isolate ligands of patched protein and assay for levels of		
CC	transcription of the patched gene. The present sequence is the mouse		
CC	patched protein.		
XX			
XX	Sequence 1434 AA;		
Query Match	9.4%; Score 646.5; DB 7; Length 1434;		
Best Local Similarity	22.9%; Pred. No. 7e-50;		
Matches	283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;		
Qy	356	LVLSPVPPVALAAGLVFELTTPDVELWSAPNSQARSEKAFDQHFQPPFTN-QVILTA	414
Db	88	LVGGLLIFGAPVLGKAAELTVEELWVEVGVSRSELNTTKIGERAMFNPQLMIQT	147
Qy	415	PNRSSRYSDLLGPKNFSGILDLDLLELLE--LQERLRHLQVWSPQAQRNISLODICY	472
Db	148	PKEEG-----ANVLITTEALQHLDSALQASRVHVVMYN-----RQWKLEHLCY	190
Qy	473	AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLLTANQTLMGQ-----	512
Db	191	KSGELITETGVMDOIIEYLYPCLIITPLDCFWEGAKLOSCTA--YLGGKPLRATWTFDPL	248

Db 191 KSGELITETGYMDQIIIEYVPCLIITPLDCLFWEKAKLQSGTA--YLLGKPPRLRWTNFDPL 248
Qy 513 -----TSQVDMKDHFLY-----CAN-----APL----- 530
Db 249 EFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPNCLNADPCDPATAPNKNSTKPLDVALV 308
Qy 531 -----TFKDG-----ALALSCMADYGAQVPPFLAIGYKKG 562
Db 309 LNGGCGQLSRKYMHWQBELIVGGTVKNATGKLSAHLQTMFLQMTPKQMYHFRGY--- 365
Qy 563 DYSEAEALITWTFSLNNYPAGDPRLAQAKLWBEAFLEMEAFQRRMAGMEQVTFTAERSLE 622
Db 366 DY-----VSHINWBE-----DRAAILEAWQTYVEVHQSVAPNSTQKVLPEFTT--TTLD 414
Qy 623 DRINKRTABDLPIFATSYIVIFLYISLALGSYSMSRVNVDKATLGLGGVAVLGAQVNA 682
Db 415 DILKSFSDVSVIRVASGYLLMLAYACLTW--LRWD--CSKSGAGVGLAGVLLVALSVAA 469
Qy 683 AMGFPSYLGIRSLVITLQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHIGRALGR 742
Db 470 GLGLCSLIGISFNAATTQVLPFALGVGVDDVFLLAHAFSETGQNKRIIPFEDRTGECLNR 529
Qy 743 VAPSMLLCSLSEAIQFELGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLDSKQ 802
Db 530 TGASVALTSISNVTAFPMALPIIPALRAFSQAAVVVFNFMVILLIPFALISMDLYNR 589
Qy 803 EASRLDVCCC-----VKPQEL-----PPQGQEGEL----- 827
Db 590 EDERLDIFCFTSPCVSRVIOEPQAYTEPHSNTRYSPPPYTSFPAETHITMQSTVQ 649
Qy 828 ----- 827
Db 650 LRTEYDPHTHYTTAEPRSEISVQPVTVTQDNLSQSPSTSTRDLSQFSDSLHCL 709
Qy 828 -----LLGFQKAYAEFLHWTIRGVVLLIFLALFGVLSYMSCHISVGLDQELAPK 879
Db 710 EPCCTKWTLSFAEKHYAFLKPKAKVGVVILLFGLLGVLSYGTTRVRDGLDLTDIVR 769
Qy 880 DSYLLDYFLNRYFEVGAQVYVTTLTGYNFSSEAGMNAICSSAGCNPFSFTQKQYAT- 938
Db 770 ETRVDYFIAAQKYFSF-----YNM-----YIVTKADYPI 801
Qy 939 -----EFPEQSYLAIPASS-----WDDFDWL-----TPSS--- 965
Db 802 QHLLYDLHKSFSNVKYMLEENKQLPQWHLFRDMLQGLQDAFSDWETGRIMPNYKN 861
Qy 966 -----CCRLYISGNKDKPCFSTVNSLNCIKNCSITMGSVRPSVEQPHKYL--PWFLN 1017
Db 862 GSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIIINPSA--FYIYLTAWVSN 915
Qy 1018 D-----RPNIKCPKGLAAY--STSVNLTSDDGOVLASRFMAYHKPKNSQD 1061
Db 916 DPVAAASQANTRPIRPEWHDK---ADYMPETRIIPAEPIEYQAQFPFYNGURDSD 972
Qy 1062 YTEALRAARELANITA--DLRKVPGTDPAPFEVPPYITTVNFYEQYLTILPEGLFMLSCL 1120
Db 973 FVEAIEKRVICNNVTSGLSSYPNG-----YPR-----LFWEQYISLHMLLSISVL 1022
Qy 1121 VTFPAVSCLLGLDLSRGLNLLSIVMILVDTVGFPMALWDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCAVFLNPNWTAGII--VMVLAMTVFELFGMGLIGIKLSAPVPPVILLIASVGIVG 1081
Qy 1181 FVSHLTRSF--AISTKPTWLERAKATISMGSAVAGVAMTNLPGTLVLGLAKAQIQLF 1238
Db 1082 FTVHVALAFLTAIGDN-----HRAMALAEHMFAPVLDG--AVSTLLGLVLMLAGSEFFIVRY 1137
Qy 1239 FFRLLMLITLLGLLHGLVLPVILSYVG--PDVNPALAEQ---KRAEAAVAAMVVASCP 1293
Db 1138 FFAVLAITLVGLNGLVLLPVLLSFFGCPCEVSPANGLNRLPTSPPEPPSVVFAVPP 1197
Qy 1294 NHPRSVS--TADNIYVNHSPSGSI-----KGAG 1319
Db 1198 GHTNNGSSDSDSESSQTTVSGISBELRYEAOQAG 1234

RESULT 51

AAR94380

ID AAR94380 standard; protein; 1434 AA.

XX AAR94380;

AC AAR94380;
DT 11-JUL-1996 (first entry)

XX Mouse patched protein.

XX Patched gene; PTC protein; embryo development; cellular regulation;
KW signal transduction; ligand; antibody; hedgehog protein.

XX Mus musculus.

XX WO9611260-A1.

XX 18-APR-1996.

XX 06-OCT-1995; 95WO-US013233.

XX 07-OCT-1994; 94US-00319745.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Scott MP, Goodrich LV, Johnson RL;

XX WPI; 1996-209842/21.

XX N-PSDB; AAT14218.

XX DNA encoding patched protein other than Drosophila patched protein - used
PT to produce antibodies which detect or inhibit patched protein ligand
PT signal transduction in cells.
PS Disclosure; Page 46-51; 70pp; English.
XX Mouse patched protein (PTC) (AAR94380) was identified as the product of a
CC cDNA clone (AAT14218) derived from mouse limb bud. It has about 38%
CC identical amino acids to Drosophila PTC. PTC protein has been proposed as
CC a receptor for hedgehog protein on basis of genetic experiments in flies.
CC Murine PTC protein can be obt'd. in large amounts by expression of the
CC cDNA clone in transformed host cells. It can be used to screen for
CC agonists and antagonists, to isolate its ligand, partic. Sonic hedgehog,
CC to assay for the transcription of ptc mRNA and to raise antibodies

SQ Sequence 1434 AA;

Query Match

Best Local Similarity 9.3%; Score 643.5; DB 2; Length 1434;

Matches 282; Conservative 172; Mismatches 420; Indels 363; Gaps 44;

Qy 356 LVLSVIPVVAALAGLVFTTETDTPVELWSAPNSQARSEKAFHQHGFPPFRFN-QVILTA 414

Db 88 LVVGLLIIFGAFVGLKAANLETNVEELWVEVGRVSRRLNVTROKIGSEAMFNPQLMIQT 147

Qy 415 PNRSSRYDSLLGLGPNFSGILDLDLLELLE--LQERLRHLQVWSPQAQRNLSLODICY 472

Db 148 PKREG-----ANVLTTBALLQHLDSALQASRVHYMYN---RQMKLEHLCY 190

Qy 473 AP-----LNPDTSLYDCINSLLOYFQNNRTLLLTANQTLMGQ----- 512

Db 191 KSGELITETGYMDQIIIEYVPCLIITPLDCLFWEKAKLQSGTA--YLLGKPPRLRWTNFDPL 248

Qy 513 -----TSQVDMKDHFLY-----CAN-----APL----- 530

Db 249 EFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPNCLNADPCDPATAPNKNSTKPLDVALV 308

Qy 531 -----TFKDG-----ALALSCMADYGAQVPPFLAIGYKKG 562

Db 309 LNGGCGQLSRKYMHWQBELIVGGTVKNATGKLSAHLQTMFLQMTPKQMYHFRGY--- 365

Query Match 9.1%; Score 628; DB 2; Length 1296;
Best Local Similarity 23.6%; Pred. No. 3.2e-48;
Matches 254; Conservative 165; Mismatches 384; Indels 272; Gaps 39;

QY 426 LLG--PRNFGILDLLLELLELRHLQVWSPAQRNIS-----LDICVAPLNLP 477
||| : : : ||| : : : : : : : : : : : :
Db 97 LLGPPLRWTFDPLEFLEELKTI----NYQVDSWEMLNKAEVGHYMDRPCINPADP 151

Query Match 9.1%; Score 628; DB 2; Length 1296;
Best Local Similarity 23.6%; Pred. No. 3.2e-48;
Matches 254; Conservative 165; Mismatches 384; Indels 272; Gaps 39;

QY 426 LLG--PRNFGILDLLLELLELRHLQVWSPAQRNIS-----LDICVAPLNLP 477
||| : : : ||| : : : : : : : : : : : :
DB 97 LLGPPLRWTFDPLEFLEELKTI----NYQVDSWEMLNKAEVGHYMDRPCINPADP 151

QY 478 DNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQ--VDWKDHFYLCANAPLPTFGQGT 536
Db 152 D-----CPATAPNKNSTKPLDMALVINGCGCHLSRKTYMHMOBELIVG--TVKNST 200
QY 537 -----ALALSCMADYCAPVPPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPLAQAKL 591
Db 201 GKLVSAHALQTFMTPQMY-----EHFKGYEY-----VSHINWNE-----DKAAAILSEA 247
QY 592 WEEAPLEEMRAFORMMAGMFQVFTTAERSLEDEINRTTAEDLPFATSVIVFIPLYISLAL 651
Db 248 WQTYVEVVHQSVAQNSTQKLSFTT--TTLDLILKSFSDSVIRVASGYLLMLAYACLTM 306
QY 652 GSYSSRSRVWDSKATLGGVAVLVGAVMAAGPFYSYLGIRSSLVILQVPPFLVLSVGA 711
Db 307 ---LRWD---CSKSGQAGVGLVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVG 361
QY 712 DNIFIEVLEYQRLPRPGPGPREVHIGRALGRVAPSMLLCSLSAICFFIGALTPMPAVRT 771
Db 362 DDVFLAHAFSETGQNKRIPEFDRTECECLKRTGASVALTISGNVTAFFMAALIPILAKA 421
QY 772 FALTGSLAVILDFLLQMSAFVALLSDSKRQBSRLDVCCC-----VKPQEL-- 818
Db 422 FSLQAAVVVVVFAMVLLIFPALLSMDLYRREDRLDIFCCFTSPCVSRVQVPEQAYTD 481
QY 819 -----PPGQGBEL----- 827
Db 482 THDNTRYSPPPYSSHSFAHETQITMQSTVOLRTEYDPHTHVYVYTTAEPRESISQPVTV 541
QY 828 -----LLGFFOKAYAPPELLHMITRGV 848
Db 542 TQDTLSQCPSTSTRDLSLQSFSSSLHCLBPCTKWTLSFAEKHYAPFLLPKPKVV 601
QY 849 VLLFLFALFGVSLYSCHISVGLDQELAPKDSYLLDYFLNRYFVEVGAQVYFVTLGY 908
Db 602 VIFLFLGLGVSIXGTRVRDGLDLTDIVPRETRVDFTAAQKYPSP-----Y 650
QY 909 NFSSEAGMNAICSSAGCNFPSTQKIQYAT-----EPPEQSYLAIPASS-----W 953
Db 651 NM-----YIVTKADYPNIQHLLYDLHRSFSNVKYVLMLENKQLPKMW 693
QY 954 VDDEFIDWL-----TPSS-----CCRLYISGPNKDKCPSVTNSL 987
Db 694 LHYFRDMLQGLQDAFSDWETGKIMPNYKNGSDGVLAYKLLVOTGSRDK--PIDISQL 751
QY 988 NCLKNCMSITGSRVPSVQFHKYL--PWFLND-----RPNKCPKGLAAY 1032
Db 752 T--KQLVDADGIINPESA--FYILTAWNSDPVAYASQANIRPHRPWVHDK--ADY 804
QY 1033 --STVNLTSDGOVLASRFMAHYKPLKNSQDYTEALRAARELAANITA--DLRKVPGTDA 1089
Db 805 MPETRLRIPAAEPFIEAQPFPFYNGLRDTSDFEAIEKVRTICSNVTSLSGLSSYPNG-- 861
QY 1090 FEVPFPYITNVFYEQVLTLPGLFMLSCLVPTFAVSCLLGLDLRSLNLLSIVMLL 1149
Db 862 ---YFP---LFEWQYIGRHWLLLFISVVLACTFLVCAVFLNPNWTAGII--VMVALMT 913
QY 1150 VDTVGFMALWDSYNAVSLINLVSAVMSVFPVSHITRSF--AISTKPTWLERAKEATIS 1207
Db 914 VELFGMMGLIGIKLSNVPVILLIASVGIVETVTHVALAFLTAISDKN--RRVALALEH 970
QY 1208 MGSAPVAGVAMTNLPGILVLGLAKAQLIOIPFRNLNLTLLGLHLGLVFLPVLSYVG-- 1266
Db 971 MPAPVLDG--AVSTLLGLVLMAGSEDFIVRYEFAVLAAILTLGLVNLGLVLLPVLSPFGP 1029
QY 1267 -PDVNPALALEQ---KRAEEAAVAVMVASCPNHPRSVSTADNIYVNHSPGSIKG 1317
Db 1030 YPEVSPANGNLNLTPTSPPEPPSVVRFAFMPGPTH--SGSDSSDSSEYSSQTTVSG 1082

RESULT 53

ADI21144

ID ADI21144 standard; protein; 967 AA.

XX

AC ADI21144;
XX 15-APR-2004 (first entry)
DE Novel human protein #119.
XX forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
XX Homo sapiens.
XX WO2003025148-A2.
XX 27-MAR-2003.
XX 19-SEP-2002; 2002WO-US029964.
XX 19-SEP-2001; 2001US-0323739P.
PR 13-SEP-2002; 2002US-00323739.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX WPI; 2003-354603/33.
DR N-PSDB; ADI21860.
XX New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX Claim 20; SEQ ID NO 395; 156pp; English.
XX The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers.
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
XX protein.
SQ Sequence 967 AA;
Query Match 8.5%; Score 588.5; DB 7; Length 967;
Best Local Similarity 21.3%; Pred. No. 9.8e-45;
Matches 216; Conservative 192; Mismatches 359; Indels 249; Gaps 37;
QY 334 HT-----LLGQFFQCGMGWTWASWPLTILVLSVIVPVALAAGVFE--TELTDPDVLWSAP 386
Db 115 HTDCLGGLSRFPQWLGMQVGAHPWIFLAPLMTALGTGLYLPKDEEDLESHYTPV 174
QY 387 NSQARSEKAFHDQHFPGPFRTNQVLTAPNRSS-----YRYSLLLGPKNPFSGIL 436
Db 175 GSPAKARRFVQGH-----FTTNDYRFSASRRSTEANFVLLVVSYSDSLDPATFAEVS 230
QY 437 DLDLLELLELLEQLERLHLOWSPAEQRNISLQDIC--YAPLNPDNTSLYDCCI--NSLLQ 492
Db 231 KLDGAVQDLRV-----AREKGSIQVQVCARYAL-----CVPPNPILY 270

Qy	493	YFQNNRTLLLTANOTLMTGQTSQVDWKDHFYCANAPLTFKDGTFALALSCWADYGA-PVF	551	OS	Homo sapiens.	
Dy	271	AWQNKTLNL-----SSISFPAY-----NHRHPLY	296	XX	Key	Location/Qualifiers
Qy	552	PFLAIGGY-----KGKDYSEARALIMTFSLNNYPAGDPLP-AQAKLWEAFLEENRAF	603	FT	Modified-site	19..22
Dy	297	LTGFFGGYGLSGMGQLLRKAMRLYYLK---TEDPEYDVQSKQWLTHLLDQFTNI	353	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Qy	604	QRMA-----GMFOYTTAERLEBIRNRTAEDLPFATSIVFLVYISL	649	FT	Modified-site	36..39
Dy	354	KNILALKKIEVGGVGLQGQEKVHVHTSLSRQLEFATSVTVPVFLAYILLILFAVT	413	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Qy	650	ALGYSYSSNRVWVDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVFPFLVLSV	709	FT	Domain	62..65
Dy	414	SCFRED-----CIRNKMCAAFGVISAFNAVVSFGLLLHIGV-PFVIIVANSFPFLILGV	467	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Qy	710	GADNIFIFVLEYQRLPRRPGREBREVHIGRALGRVAPSMLLCSLEAICFFLIGALTPMPAV	769	FT	Modified-site	79..82
Dy	468	GVDDMFIMISAWHK-TNLAGDIRE-RMGNVYSKAAVSITITITNIALYTGIMSPRSV	525	FT	Modified-site	166..952
Qy	770	RTFALTSLGLAVLDFLLQMSAFVALLSLDSKROEASRLDVCCKVQPELP-----	819	FT	Modified-site	/note= "Patched domain"
Dy	526	QCFCIYTGTLTLCFYNYITCEGAFNALDGREV-----VLCMLKKADPKWSPFKKPC	580	FT	Modified-site	190..193
Qy	820	-----PQGQSEGL--LLGFFQKAYAPFLHMTIRGVVLLFLALFGVLSYMSCHISVGL	871	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Dy	581	PPFGSVPDHGTDIHPMSLFFRDYFGPFLTRSESKYFVVFVLYIISIIYGCFFHVSGL	640	FT	Modified-site	192..195
Qy	872	DEGLALPKDSYLLDFELFNRYF-EVGAPVYPTVLTYNFSSEAGMNAICSSAGCWNFSF	930	FT	Modified-site	/note= "N-glycosylation site"
Dy	641	DLRNLASDDSYITPYFNVENEFSDYGRVMVIVTKVDYMDK-----DV	685	FT	Modified-site	194..196
Qy	931	TKIQVATEFPQSQSLAIP--ASSWVDDFDLMTLTPSSCCRLYISGNPKDKFCPTVNSLN	988	FT	Modified-site	200..202
Dy	686	RKLENCCKIFKNVYVDKNTFELWDVYQ-----YLGKNSQD---PNEKNT--	730	FT	Modified-site	/note= "Protein kinase C phosphorylation site"
Qy	989	CLKNCSITMGSRVPSVEQHKYLPWFLNDRNIKCPKGLAAYSTSVNLSDGQVLASR	1048	FT	Modified-site	201..204
Dy	731	-----FMNIPDLSNFPN-----FQHDINISSNBIISR	761	FT	Modified-site	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
Qy	1049	FMAYHKPLKNSQDYTEALBAARELANITADLRKVPGTDPAPFVPTTITNVEYQYLTI	1108	FT	Modified-site	215..218
Dy	762	GFITQTTDVSSS-----AKKXI---LLFQURRI-AEDCQPLMVYNOAFYFQYAAI	809	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Qy	1109	LPB-----GLPMLSICLVP-----TFVAVSCILLGLDLRLSGLLNLISVILV	1150	FT	Modified-site	219..222
Dy	810	LEDTVNLVLSAAMFVLSLLIPYPLCSLWVTFAIGSVIVGV-----	852	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Qy	1151	DTVGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLERAKEATISMGS	1210	FT	Modified-site	225..228
Dy	853	--TGFMFAFKVNLDSISMLNLVICIGSFDFSAAHSYAFVSSSQSVNOKSVEALVLLGY	910	FT	Modified-site	/note= "Casein kinase II phosphorylation site"
Qy	1211	AVFAGVAMTNLPGLVILGLAKAOLIOIFPFLNLLITLGLLHGLVFLPVLVSYVG	1266	FT	Modified-site	230..233
Dy	911	PVLQS-AISTIIIGVCVLAAKAYIFRT-PPKIMFLVMIFGAAGLLIFIPVLITFFG	964	FT	Modified-site	275..278
RESULT 54					383..540	/note= "Tyrosine kinase phosphorylation site"
ABB98519					418..422	/note= "Sterol-sensing domain, SSD, domain"
ID ABB98519 standard; protein; 954 AA.					504..509	/note= "N-myristoylation site"
XX					508..510	/note= "N-myristoylation site"
AC					530..533	/note= "Protein kinase C phosphorylation site"
XX					535..540	/note= "N-glycosylation site"
XX					541..544	/note= "N-myristoylation site"
XX					561..563	/note= "Amidation site"
XX					572..575	/note= "Protein kinase C phosphorylation site"
XX					597..600	/note= "Casein kinase II phosphorylation site"
XX					662..664	/note= "Casein kinase II phosphorylation site"
XX					678..681	/note= "Protein kinase C phosphorylation site"
XX					681..688	/note= "N-glycosylation site"
XX					692..695	/note= "Tyrosine kinase phosphorylation site"
XX					737..740	/note= "N-glycosylation site"
XX					740..743	/note= "N-glycosylation site"
XX					746..748	/note= "Casein kinase II phosphorylation site"

KW	Huntington's disease; drug addiction; bone disease; skin disease; ulcer;
KW	infertility; lung disease; pancreatic disorder; diabetes; osteoporosis;
KW	therapy.
XX	Homo sapiens.
XX	WO9953058-A1.
PN	21-OCT-1999.
XX	
PD	02-APR-1999; 99WO-US007417.
XX	
PF	15-APR-1998; 98US-00060939.
XX	
PR	(GETH) GENENTECH INC.
XX	
PA	De Sauvage FJ, Carpenter DA;
XX	
PI	WPI; 1999-620428/53.
XX	
DR	N-PSDB; AA231717.
XX	
PT	New isolated human patched-2 gene, used to develop products for treating,
XX	e.g. cancer and Alzheimer's disease.
XX	Claim 14; Fig 1; 124pp; English.
PS	
XX	This sequence represents the human patched-2 (ptch-2) protein of the
CC	invention. The patched-2 polypeptides are signalling molecules,
CC	specifically for signalling and mediator molecules in the hedgehog (hh)
CC	cascade which are involved in cell proliferation and differentiation.
CC	They can be used for the treatment of disorders which are mediated at
CC	least in part by hh, especially Dhh, e.g. testicular cancer. They can
CC	also be used for treating degenerative disorders of the nervous system,
CC	e.g. Parkinson's disease, memory deficits, Alzheimer's disease, Lou
CC	Gehrig's disease, Huntington's disease, schizophrenia, stroke and drug
CC	addiction. Patched-2 agonists can be used to treat gut diseases, bone
CC	diseases, skin diseases, diseases of the testis (including infertility),
CC	ulcers, lung diseases, diseases of the pancreas, diabetes, and
CC	osteoporosis. Antagonists or agonists of patched-2 may be used for
CC	treating disorders or creating a desirable physiological condition
CC	effected by blocking hh signalling, especially Dhh signalling, e.g.
CC	contraction or infertility treatment. The products can also be used for
CC	detection, diagnosis, drug screening and production of transgenic animals
XX	
XX	Sequence 1203 AA;
Query Match	8.5%; Score 584; DB 2; Length 1203;
Best Local Similarity	22.9%; Pred. No. 3.8e-44;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;	
QY	306 RVAPARDKSNVDPKKGTSLSDKL--SFSTHTLLGQFFQG---WGTWASWPLTILVLS 359
DB	3 RSPPLRELPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLLPSLGGCIQRHCGKVLFLG 62
QY	360 VIPVVALAGLVELTITDPELWSAPNSQARSEKAFHQHFG--PFFRTNOVILTAPNR 417
DB	63 LLAAGFALGLRMAIETNLEQLWVEGSRVSOELHYTEKUGEEAAIYSQMLIQARQE 122
QY	418 SSVRYDLSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNLSQDICY---AP 474
DB	123 GEN-----ILTPALG-----LHLQAALTASKVQVSLYKGSWDLNKICYKGVF 166
QY	475 LNPDN-----TSLYDCINSLLYQFQNNRTLLLTANQTLMGQTQVDWKD----- 520
DB	167 LIENGMIETWIEKLFPFCVILTPLDCEWEGAK---LQGGSAFLPCRPDIQWNTLDPQLLE 223
QY	521 -----HFLYCANAPLTFKDGTAALASQADYGA 548
DB	224 ELGPFASLEGFRELLDKAQGVQAVGRPCRLHPDDLHCPPSPAPNHHSEQAENVAHELSGGC 283
QY	549 PVFPF-----LAIGGYKGDYSE---AALIMTF-----SLNNYPAGDPRLAQAKL 591
DB	284 HGFSGKFMHMQEELLGGMARDPQGLLRARALQSTFLLMSPRQLYEHFRGDYQTHDIGW 343
QY	592 WEAFLEEMRAFORMAGMFQVITFAERSLE-DEINRTTAEDLPIFATS-----Y 640
DB	344 SEEQASTVLQAWQRRFVQLAQEALPENASQOIHAFSSTLUDDI-LHAFSEVSARVVGGY 402
QY	641 IVIFLYISIALGSSYSRSMVVDKATLGLGGVAVVLGAVMAAMGFYSYLGIRSSVLILQ 700
DB	403 LMLAVACTM---LRWD--CAQSQSGVLGAVLLVALAVASGLGLCALLGITFNAATQ 457
QY	701 VVPFLVSGADNIFIVLEYQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEACFFL 760
DB	458 VLPFLALGIGVDDVFL--LAHAFTALPGTFLQERMGECLQRTGTSVVLTSINMAAPLM 515
QY	761 GALTPEAVTEALTSGLAVIDFLQMSAFVALLSLDKRQASRLDVCVCC----- 812
DB	516 AALVPIPALRAFLQAIAIVGCTFVAVMLVFPALLSLDLRRRCORLDVLCFSSPCSAQ 575
QY	813 ---VKQDEL-----PPGQGEGL-- 827
DB	576 VIQLPQELGDGTVPGVIAHLTATVQAPTHCEASSQHWVITLPQAHLPVPPSPDLGSEL 635
QY	828 -----LLG-----PFQAYAPELLHWTIRGVVLLFLA 855
DB	636 FSPGGSTRDLLQGEETRQKAACKSLPCARWNLAHFARYOFAPLLOSHAKAIVLVLFGA 695
QY	856 LFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPVYFVTTLGYNFS-SEA 914
DB	696 LLGLSLYGTATVQGLLALTVDVPRGTEHAFLSQAQLAYFSL-YEVALVTQGGFDYAHSQ 754
QY	915 GMAICSSAGCNNSFTQKIQYATEFPFQSYLAIPAS-----SWDDFIDMTTPSSCC--- 967
DB	755 AL-----FDLHQRFSSL-----KAVLPPATQAPRTWLHYRNMLQGIQAAPDQ 798
QY	968 -----RUYISGPNKDKPCPTVNSLNCNMSITWGSVPSVQEHKYLFWFLND 1018
DB	799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQEPLD-FSQLTRKLVD 846
QY	1019 RPNIKCPKGGGLAAVSTSVNLTSDQVLASRPMAYHKP----- 1055
DB	847 REGLIPP--LFYMLTIVWSDDLGLAASQANFYPPPEWLHKYDTTGENLRIPPAQP 904
QY	1056 -----LKNQSDYTLAARAELAANI-TADLRKVPGTDPAPFVPTTINVPY 1102
DB	905 LEFAQPELLRGLQKTADFEAIEGARAACAEAGQAGVHAYPSGSPF-----LFW 954
QY	1103 EQLTILPEGLFMLSIC--LVPTFAVSCLLGLDLRSGLLNLSTVILVDTVGFMAIWD 1160
DB	955 EQLYGL--RRCFLLAVCILLVCTFLVCAILLNLPWTAGLI-VLVLAMTVELFGIMGFLG 1011
QY	1161 ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKBATISMGSAVFAGVAMTN 1220
DB	1012 IKLSAIPVVLVASVIGIVEFTVHVALGF-LTTQGSRLRAAHAALEHTFAPVTDG-AIST 1069
QY	1221 LFGILVGLAKAQLIOIPFRNLNLLITLLGLHLGVFLPVILSYVGPVDPVNPALAEQKA 1280
DB	1070 LLGLMLAGSHPDFIVRYFFAALTVLTLGLHLGVLLPVLISILGPP--PEVIQMYKES 1127
QY	1281 EBAVAV-----NVASCNHPRSRVSTADNIYVN 1308
DB	1128 PEILSPAPPQGGGLRWGASSLPSQSFARVTTSMVTVAIH 1165
RESULT 56	
AAE19829	
ID	AAE19829 standard; protein; 1203 AA.
AC	AAE19829;
XX	
DT	18-JUN-2002 (first entry)
XX	
DE	Human patched-2 (Ptch-2) protein.
XX	
KW	Human; patched-2; Ptch-2; cell proliferation; differentiation; therapy;

cytostatic; testicular cancer; hedgehog protein signalling.

Homo sapiens.

Key Location/Qualifiers

Domain 57..76 /label= TM1 /note= "Transmembrane domain" 394..414 /label= TM2 /note= "Transmembrane domain" 427..449 /label= TM3 /note= "Transmembrane domain" 458..477 /label= TM4 /note= "Transmembrane domain" 502..524 /label= TM5 /note= "Transmembrane domain" 532..553 /label= TM6 /note= "Transmembrane domain" 687..705 /label= TM7 /note= "Transmembrane domain" 964..984 /label= TM8 /note= "Transmembrane domain" 988..1010 /label= TM9 /note= "Transmembrane domain" 1014..1033 /label= TM10 /note= "Transmembrane domain" 1058..1077 /label= TM11 /note= "Transmembrane domain" 1091..1114 /label= TM12 /note= "Transmembrane domain"

US6348575-B1.

19-FEB-2002.

15-APR-1999; 99US-00293505.

15-APR-1999; 98US-0081884P.

(GETH) GENENTECH INC.

De Sauvage F, Carpenter DA;

WPI; 2002-215260/27.

N-PSDB; AAD31576.

Native human patched-2 polypeptide for treating disorders caused by Hedgehog protein signalling such as testicular cancer, and for screening cDNA libraries.

Example 1; Fig 1; 82pp; English.

The invention relates to an isolated sequence comprising a native human patched-2 (Ptch-2) polypeptide. The invention also relates to signalling and mediator molecules in the hedgehog (hh) cascade which are involved in cell proliferation and differentiation. The isolated sequence is useful for the treatment of disorders which are linked to Hedgehog, especially Desert hedgehog expression, such as testicular cancer. It may also be used as a hybridisation probe in a cDNA library to isolate Ptch-2 or its homologues, and to diagnose whether a disorder is driven by Ptch-2 or Hedgehog protein signalling. The present sequence is human patched-2 (Ptch-2) protein

XX	SQ	Sequence 1203 AA;	Query Match
XX			Best Local Similarity 22.9%; Pred. No. 3.8e-44; Length 1203;
XX			Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
QY	306	RVPARDKSKMVDPKKGTSLSDKL--SFSTHTLLGQFFQG---WGTWVASWPLTTLVLVS	359
DB	3	RSFPLRELPPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLLPSLGGCQIRHCGKVLFLG	62
QY	360	VIPVVALAAGLVFTLTDVPELWAPNSQARSEKAFHDQHF--PFFPTNVOILTAPNR	417
DB	63	LLAFGALALGLRMAIETNLEQLWVEGSRVSEQLHYTKELGEEAAVTSQMLIQTAQRE	122
QY	418	SSRYDSSLGLGPKNFSGILDLLLELELEQLERHLQVMSPEAQRNISLDQICY---AP	474
DB	123	GEN-----ILTPEALG-----LHLQAALTASKVQVSLYKSWDLNKKICYKSGVP	166
QY	475	LMPDN-----TSLYDCCINSLLQYFQNNRTILLLTANQTLMGQTSQVDWKD-----	520
DB	167	LIENGMIEMIEKLPFCVILITPLDCFWEGAK---LQGSAYLPGRPDIQNTNLDPEQLLE	223
QY	521	-----HFLYCANAPLTFKDGTTALALSCMADYGA	548
DB	224	ELGPPASLEGFRELDDKAQVQAYVGRPCLPDHLPPDLHCPPSAPNHHSRQAPNVAHELSGGC	283
QY	549	PVFPF-----LAIGGVKGYSE---AREALIMTF-----SLNNYPAGDPRLAQAKL	591
DB	284	HGFSHKFMHQBELILGGWARDPOGBELLRAELAQSTFLLMSPRQLYEHPRGDIYQTHDIGM	343
QY	592	WEAEFLBEMRAFQRRMAGMFQVTFTAERSLE--DEINRTTAEDLPFATS-----Y	640
DB	344	SEEQASTVLQAWQRFVQLAQEALPENASQIHAFSSTLLDI--LHAISEVSAARVVGYY	402
QY	641	IVIFLYISLALGSYSSWSRVMDSKATIGLGVAIVLGAVMAAMGFFSYLGRSSLVILQ	700
DB	403	LLMLAYACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGICALLGITFNAATQ	457
QY	701	VVPELTVSGADNIFIVLEYQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEACIFPL	760
DB	458	VLPFLALGIGVDDVFL--LAHAFTALPGTFLQERMGECLQRTGTSVLVLTINNMAAFILM	515
QY	761	GALTPMAVTRTALTSLGLAVILDFLLOMSAFVALISLDSKROEASRLDVCCC-----	812
DB	516	AALVPIPALRAFSLQAAIVGCTFVAVMLVFPAILSLDLRRHRCQLDLVLCFSSPCSAQ	575
QY	813	---VKPQEL-----PFGQEGEL--827	
DB	576	VIQLPQELGDTVPVGIHLTATVQAFTHCEASSQHVVTILPQAHLVPPSPDPLGSEL	635
QY	828	-----LLG-----FPQKAYAPFLHWTGCVLLPLA	855
DB	636	FSPGGSTRDLLGQBEETROKAAKSLPCARNLAHFARYQFAPLQLQSHAKAIVLVFGA	695
QY	856	LFGVLSYCMCHISVGLDQELALPKDSYLLDFLNRIFYEVGAPVYFVTTLGYNFS--SEA	914
DB	696	LLGLSLYCATLVQDGLATDVVPGTKEHAFSLAQRLYFSL--YFVALVTQGGFDYAHSGR	754
QY	915	GMNAICSSAGCNFSTQKIQYATEFPQSYLAIPAS-----SWVDDFDIMLTPSSCC---	967
DB	755	AL-----PDLHQRFSSL-----KAVLPPPTQAPRTWLHYRNWLQIGIQAAAPDQ	798
QY	968	-----RLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVQFPHKYLFWFLND	1018
DB	799	DWASGRITRHSYRNG--SED-----GALAYKLLIQTDGAQEPFLD--FSQLATTKRLVD	846
QY	1019	RPNIKCPKGGGLAAVSTSVNLTSQGVLASRFMAVHKP-----	1055
DB	847	REGLIPPE--LIFYMGLTWVSSDPLGLAASQANFYPPPPPEWLHKYDVTGEMLRIPPAQP	904
QY	1056	-----LKNSQDYTBALRAARELANI--TADLRKVPGTDPAPFVFPYTTITNVFY	1102

Db 905 LEFAQFPFLRLGLQKTADFEAIEGARRACACAGQAGVHAYPSGSPF-----LFW 954
Qy 1103 EQYLITPEGLFMLSIC--LVPTFAVSCLLGLDLRGLMLLSIVMLVDTVGFMALWD 1160
Db 955 EQYGLL--RRCFLAVCLLAVCTFLVCALLLNPNWAGLI-VLVLAAMTWELFGIMGFLG 1011
Qy 1161 ISYNAVSLINLVASGVMSVEFVSHITRSFALSTKPTWLERAKEATISNGSAVFAGVAMTN 1220
Db 1012 IKLSAIPVVILVASVIGVETVHVALGF-LTTQGSRLNRAHAHEHTFAPVTDG-AIST 1069
Qy 1221 LPGAIVLGLAKAQLQIIFERLNLITLLGLLHGLVFLPVLISVVGPDVNPALALEQKRA 1280
Db 1070 LGLLMLAGSHPDFIVRFFAALTVLLGLLHGLVLLPVLISLILGPP-PEVIQMYKES 1127
Qy 1281 BEAFAV-----MVASCPNHPRSRVSTADNIYN 1308
Db 1128 PILSPAPQGGGLRWAGSSLPQSFARVTTSMVTVAIH 1165
AC ABG74104 standard; protein; 1203 AA.
XX ABG74104;
XX
DT 02-APR-2003 (first entry)
XX Human patched-2 protein.
XX
KW Human; patched-2; Dhh signalling; proliferation; differentiation;
KW Chromosome 1p33-34; male contraceptive; infertility; spermatogenesis;
KW basal cell carcinoma; neurodegenerative disorder; memory deficit;
KW Alzheimer's disease; Parkinson's disease; Lou Gehrig's disease;
KW Huntington's disease; schizophrenia; stroke; drug addiction; gut disease;
KW bone disease; skin disease; testicular disease; ulcer; lung disease;
KW pancreatic disease; diabetes; osteoporosis; desert hedgehog.
XX
OS Homo sapiens.
XX
XX US2002156245-A1.
XX
XX 24-OCT-2002.
XX
XX 20-NOV-2001; 2001US-00990046.
XX
XX 15-APR-1998; 98US-0081884P.
XX 15-APR-1999; 99US-00293505.
XX
XX (GETH) GENENTECH INC.
XX
XX De Sauvage FJ, Carpenter DA;
XX WPI; 2003-182650/18.
XX N-PSDB; ABX15919.
XX
XX New nucleic acid, useful for manufacturing a medicament for diagnosing or
XX treating a disorder that is modulated by Desert hedgehog (Dhh) signaling
XX e.g., testicular cancer.
XX
XX Claim 14; Fig 1; 85pp; English.
XX
XX The invention relates to a new isolated nucleic acid encoding a
XX polypeptide having patched-2 biological activity, comprises DNA having at
XX least 95% sequence identity with a DNA molecule or its complement
XX encoding: (a) a human patched-2 polypeptide comprising the sequence
XX 'ABG74104'; or (b) the same mature polypeptide encoded by the cDNA (ATCC
XX Deposit No. 209778 designation). Also included are a vector comprising
XX the nucleic acid, a host cell transformed with the vector, a process for
XX producing patched-2 polypeptides, an isolated native sequence of human
XX patched-2 polypeptide, a chimeric molecule comprising the vertebrate
XX patched-2 polypeptide and a heterologous (aa) sequence, an antagonist of
XX patched-2 (that blocks, prevents, inhibits and/or neutralises the Desert
XX hedgehog (Dhh) function in the Dhh signalling pathway), an agonist of

CC patched-2 that stimulates or enhances the normal functioning of patched-2
CC in the Dhh signalling pathway, screening for ant/agonists of patched-2
CC and diagnosing to determine whether a particular disorder is modulated
CC The nucleic acid is useful for manufacturing a medicament for diagnosing
CC or treating a disorder that is modulated by Dhh signalling e.g. tumour,
CC basal cell carcinoma, neurodegenerative disorders, memory deficit,
CC Huntington's disease, Parkinson's disease, stroke, drug addiction, gut
CC diseases, bone diseases, skin diseases, testicular diseases, infertility,
CC diseases, pancreatic diseases, diabetes, osteoporosis and infertility.
CC Patched-2 may also be used to develop male contraceptives. The gene for
CC Patched-2 is located on human chromosome 1p33-34. The present sequence
CC represents human patched-2 protein
XX
XX Sequence 1203 AA;
Qy Query Match 8.5%; Score 584; DB 6; Length 1203;
Best Local Similarity 22.9%; Pred. No. 3.8e-44;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
Qy 306 RVAPARDKSKMVDPKKGTSLSDKL--SFSTHTLLGQFFQG---WGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLWLRAYFQSLGLGCIQRHCKVFLFG 62
Qy 360 VIPVVALAAGLVFTELTDPVELWSPNSQARSEKAFHDQHFQ--PFRNTNQVLLTAPNR 417
Db 63 LLAFGALALGLRMAIETNLEQLWVEGVSQBELHYTKELGAEAAYSQMLIQTARQE 122
Qy 418 SSRYDLSLLGPKNFSGILDLLELLELEQLERHLQVMSPEAQRNLSQDICY---AP 474
Db 123 GEN-----ILTPEALG-----LHQAALTASKVQSVLSYKSWDLNKKICYKSGVP 166
Qy 475 LNPDN-----TSLYDCCINSLLQYFQNNRTLLLTANTQMLGQTSQVDKX----- 520
Db 167 LIENGMEWMEIKLPPPCVILLTDLPCFWEKAK---LQGSAYLPGRPDIQNTLNDPEQLLE 223
Qy 521 -----HFLYCANAPLTFKDGTTALALSCWADYGA 548
Db 224 ELGPPASLEGPRELLDKAQVQAVYGRPCLPDLDLHCPSPAPNHHSHROAPNAHELSGCC 283
Qy 549 PVFPF-----LAIGYKGDYSE---AEALITWP-----SLNNYPAGDPRLAQAKL 591
Db 284 HGFSHKFMHQBELLLGGWARDPQGLLRALAGSTFTLLMSPRQLYEHFRGDYQTHDIGW 343
Qy 592 WEEAFLEEMRAFORRMAGMFQVTTAERSLE-DEINRTTABEDLPFAFS-----Y 640
Db 344 SEEQASTVLQAWQRFFVLAQEALENASQIHAFSSITLDDI-LHAFSEVSAARVWGY 402
Qy 641 IVIFLYISLALGSYSSWSRVWDSKATILGLGVAVVLGAVMAAMGFFSYLGRSSILVILQ 700
Db 403 LILMAYACVTM---LRWD---CAQSQSVGLAGVLVALAVASGLGLCALLGITFNAATTQ 457
Qy 701 VVPLVLVSGADNIFVLEVQRLPRRPEPEVHIGRALGRVAPSMILCSLEAICFFPL 760
Db 458 VLFFLALGIGVDVDFL--LAHFTALPQTPLOERMGECLQRTGTSVLTINNMAAFILM 515
Qy 761 GALTTPMVAVRTPALSGLAVIDLDFLLQMSAFVALLSLDSKQESARLDVCC----- 812
Db 516 ALVLPALRAFSLQAQAIWVGCTFFVAVMLVFPAILSLDLRRRCORLDLVCCFSFSCSAQ 575
Qy 813 ---VKPQEL-----PPQGGSGEL-- 827
Db 576 VIQLLPQELGDTGTVFVGIHAHLTATVQAFTHCBASSQHVVTILPPQAHLPVPPSDPLGSEL 635
Qy 828 -----LIG-----PFQAYAFPLHMTIRGVVLLFLFA 855
Db 636 FPGGSTRLLQGEETROKAKCSLPCARWNLAFARVQFAPLQLQSHAKAIVLVFGA 695
Qy 856 LFGVSLYSCHISVGLDQDLALPKOSYLDYFLFNRYFEVGAIVYFVTTLGNFS-SEA 914
Db 696 LLGLSLYGATLVQDGLATDVVPRGPKHAFLSAQLRYFSL-YEVALVTQGGFDYAHQSR 754
Qy 915 GMAICSSAGCWNFSFTQKIQATPEPQSYLAIPAS-----SWDDFDWLTPSSCC--- 967

870 -GLAASQANFYPPPEWLHDKYDTTGENLRIPAAQPLEFAQFPFLLHGLQKTADFVEAIE 928
1068 AARELAANI-TADLRKVGCTDPAFEVPPYTTITNVFYEQYLTILPEGLFMLSIC--LVPTF 1124
929 GARAACTEAGQGVHAYPSGSPF-----LFWEQYLG--RRCFLLAVCILLVCTF 976
1125 AVSCLLGLDLSRGLNLLSVMILVDTVGPMALWDISYNVAVSLINIVSAGMSVEFVSH 1184
977 LVCALLLSPWTAGLI-VLVLAAMTVELFGIMGFLGKLSAIPVILVASIGIGVEFTVH 1035
1185 ITRSPAISTKPTWLERAKEATISGSAVAGV---AMTNLPGILVIGLAKAQLQIIFER 1241
1036 VALGFLTSHGSRNLRAA-----SALEQTFAPVTDGAVSTLLGLLMLAGSNPDFIIRFFV 1090
1242 LNLTLTLLGLLGLVFLPVILSYGVP 1267
1091 VLTUTLTLGLLGLLGLPVILSYLGP 1116

RESULT 59
AAY92703
ID AAY92703 standard; protein; 1203 AA.
XX
AC AAY92703;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human patched 2 (PTCH2) protein splice variant.
XX
XX Patched 2; PTCH2; lp32-35; tumour suppressor; familial melanoma CMW1;
KW familial adenomatous polyposis; hMomi; Michelin Tire Baby Syndrome;
KW sonic hedgehog; SHH; development; carcinogenesis; cytostatic;
KW basal cell carcinoma; gene therapy; chromosome lp32-35; splice variant.
XX
OS Homo sapiens.
XX
XX WO200020037-A1.
XX
XX 13-APR-2000. 99WO-SE001784.
XX
XX 06-OCT-1999; 98SE-00003393.
XX
XX 06-OCT-1998; 98SE-00003393.
XX
XX (KARO-) KAROLINSKA INNOVATIONS AB.
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Zaphiropoulos PG, Unden AB, Toftgard R, Rahnama F;
PI Hollingsworth RE;
XX
XX WPI; 2000-303645/26.
XX
XX N-PSDB; AAA09084.
XX
XX Isolated human protein capable of participating in human patched
PT Gene/Sonic hedgehog pathway during embryonic development is used in
PT medicament for treatment of condition involving tumors such as basal cell
PT carcinoma.
XX
XX Disclosure; Page; 55pp; English.
XX
XX This protein is a splice variant encoded by the novel human patched 2
CC gene (PTCH2), which has been localised by radiation hybrid mapping to
CC chromosome lp32-35 with DIS211 and W11404 as closest flanking markers and
CC with an estimated localisation 5.5 cR from DIS443. This region is often
CC lost by LOH in various different tumour types, such as neuroblastoma,
CC melanoma and breast and colon cancer. PTCH2 is a candidate for a tumour
CC suppressor gene in this region. It is also a candidate gene for
CC involvement in familial melanoma CMW1, modifier locus for familial
CC adenomatous polyposis hMomi and Michelin Tire Baby Syndrome. PTCH2 is
CC capable of participating in the human patched gene/sonic hedgehog
CC (PTCH/SHH) pathway during embryonic development and/or carcinogenesis.
CC The isolated human protein is useful as a medicament for the treatment of
CC a condition involving tumours such as BCC (basal cell carcinoma). The

CC nucleic acid is useful in gene therapy, and for use as a probe, primer or
CC a diagnostic agent. Note: This sequence was constructed using information
CC from Figure 2B and the protein sequence given on pages 41-44 of the
CC specification
XX
SQ Sequence 1203 AA;
Query Match 8.4%; Score 583; DB 3; Length 1203;
Best Local Similarity 22.9%; Pred. No. 4.7e-44;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLGGFFQG----WGTWVASWPLTILVLS 359
3 RSPFLRELPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLFLSLGCGIQRHCGKVLFLG 62
360 VIPVVALAGLVTELTDPVELWSAPNSQARSEKAFHQHFG--PFERNNOVILTAPNR 417
63 LLAFGALALGLRMATIIETNLEQLWVEVSGRSVSELHYTKEKLEEAAYTSQMLIQIARQE 122
418 SSYRYDSLILGPKNFGILDLLELLELLEQLRHLQVMSPEAQRNLSLDICY---AP 474
123 GEN-----ILTPEALG-----LHLQALTASKVQVSLYKGSWDLNKICYKSGVP 166
475 LNPDN-----TSLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDWKD----- 520
167 LIENGMIERMEKLPFCVILTPLDCFEWGAKE---LOGGSAYLFGRPDIQNTLDPQOLLE 223
521 -----HFLYCANAPLTFKDGCTALALSCMADYGA 548
224 ELGPFASLEGFRELDDKAQVQAVVGRPCLDHPPDPPAPNHHSRQAPNVAHELSGGC 283
549 PVRFPP-----LAIGGYKGYSE---AEALIMTF-----SLNNYPAGDPRLAQAQL 591
284 HGFHKKFMHQEELLGGMARDPQGLLRARALQSTFLLMSPRQLYEHFRGDYQTHDIGM 343
592 WEAFLEEMRAFQRMAGMFQVTTAERSLE-DEINRTAEDLPIFATS-----Y 640
344 SEQASTVLQAWQRRFVQLAQEALPENASQIHFSSSTLDDI-LHAFSEVSAARVVGYY 402
641 IVIFLYISLALGSSYSSSRVWVDSKATLGLGGVAVVLGAVNMAAMGFFSYLGIRSLVLQ 700
403 LMLAYACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGCLALLGITFNAATQ 457
701 VVPFLVLSVGADNIFIVLEYQRLPRRPGEPREHIGRALGRVAPSPMLCSLSEACIFPL 760
458 VLPFLALGIGVDDVFL--LAHAFTEALPGTFLQERMGECLOQTGTSTVVLTSINNMAAFIM 515
761 GALTMPAVRTFALTSLGLAVILDFLQMSAFVALLSLDSKQESRLDVCC----- 812
516 AALVPIPALRAFSLOAAIIVGCTFVAVMLVFPAILSLDLRRHRCORLDVLCFSSPCSAQ 575
813 ---VKPQEL-----PPPGQEGEL-- 827
576 VTQIILPQELGDGTVPVGIHLTATVQAFTHCEASSQHVVTILPQAHLVPPSPDPLGSEL 635
828 -----LLG-----PFQKAYAPFLHMTITRGVLLLLPLA 855
636 PSPGGSTRDLGQEBEETROKAAKSLPCARNWLAHFARYQFAPLQLQSHAKAHLVILVFGA 695
856 LFGVSLYSMCHLSVGLDQELALPKDSYLLDYFLFLNRVPEVGPVVFVTTGLGVNFS-SEA 914
696 LLGLSLYGATLVQDGLALTVDVVPRTGKEHAFLSAQLRYFSL-YEVALVTQGGFDYAHQR 754
915 GMAICSSAGCNFNFSFTQIKIYATFEPPEQSYLAIPAS-----SWVDDFDWLTPSSCC--- 967
755 AL-----FDLHQRFSSL-----KAVLPPPTATQAPRTWLHYRYNWLQIQAAFDQ 798
968 -----RLYISGNKDKFCPSTVNSLNCNKMSITWGSVRPSVEQFHKYLPWFLND 1018
799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQELLD--FSQUTTRKLVLD 846
1019 RPNIKCPKGLAAYSTVNLTSQGVQLASRFMAYHKP----- 1055

Db 847 REGLIPE--LFVMTGLTVWSSDPLGLAASQANFYPPPEWLHDKYDVTGTFNRIIPPAQ 904
 QY 1056 -----LKNSQDYTEALRAARELAANI-TADLRKVPCTDPAFEVFPVTIINVEY 1102
 Db 905 LBEAQPFLRLGLOKTADFVEALEGARACAEAGQAGVHAYPSGSP-----LFW 954
 QY 1103 EOYLITLPLGLFMLSIC--LVPTFAVSCILLGLDRLSGLLNLSIVMILVDTVGFNMLWD 1160
 Db 955 EOYGLG--RRCFLAVCILLVCTFLVCALLLNPNWTAGLI-VLVLAMTVVELFGINGFLG 1011
 QY 1161 ISNAVSLNLSAVGMSVEFVSHITRSPAISTKPTMLERAKEATISMGSAVPAGVAMTN 1220
 Db 1012 IKLSAIPVILVASGVIGFVTHVALGP-LITQSGNRNRAHALEHTPVTG-AIST 1069
 QY 1221 LFCILVLGLAKAQLOIIFFRNLNLITLGLLHGLVFLFVILSVYGVDPNPALAEQKRA 1280
 Db 1070 LLGLMLAGSHDFIVRYFFAALTVLTLGLLHGLVLLPVLLSILGPP--PEVIQWKES 1127
 QY 1281 EBAVAV-----WVASCNHPHSRSTADNIYVN 1308
 Db 1128 PEILSPPAPQGGRLWGAASSLPQSPARVTTSMTVAIH 1165
 RESULT 60
 AAY28444
 ID AAY28444 standard; protein; 1203 AA.
 AC AAY28444;
 DT 03-DEC-1999 (first entry)
 XX Human ptc-2 protein.
 KW Patched-2; ptc-2; human; hedgehog receptor; nootropic; neuroprotective;
 KW antiinflammatory; antiparkinsonian; cardiant; antiarthritic; screening;
 KW modulator; antagonist; agonist; cellular proliferation; neuronal tissue;
 KW testicular tissue; osteogenic tissue; chondrogenic tissue; disease;
 KW graft; transplant; treatment; nervous system injury; chemical injury;
 KW vasa; infection; inflammatory; tumor-induced injury; ageing;
 KW Alzheimer's disease; chronic neurodegenerative disease; innervation;
 KW Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;
 KW spinocerebellar degeneration; multiple sclerosis; autonomic disorders;
 KW peripheral nervous system; smooth muscle; endocrine tissue; tachycardia;
 KW atrial cardiac arrhythmia; cell differentiation; chronic pain syndrome;
 KW lesion-induced death; neuron regeneration; damage repair; skeletal;
 KW cartilage; osteogenesis; arthritis; bone fracture; hereditary disease;
 KW prosthetic cartilage device; spermatogenesis; fertility enhancer.
 XX Homo sapiens.
 OS
 XX WO9929854-A1.
 PN
 XX 17-JUN-1999.
 PD
 XX 08-DEC-1998; 98WO-05026009.
 PF
 XX 08-DEC-1997; 97US-0067940P.
 PR
 XX (ONTO-) ONTOGENY INC.
 PA
 XX Bumcrot DA;
 PI
 XX WPI; 1999-561298/47.
 DR
 XX N-PSDB; AAX89478.
 DR
 XX New human patched-2 (ptc-2) genes and proteins, useful in the treatment,
 PT prevention and/or reduction of the severity of neurological conditions.
 PT
 XX Claim 3; Page 73-77; 80pp; English.
 PS
 XX This invention describes a novel recombinantly produced human patched-2
 CC (ptc-2) polypeptide which has nootropic, neuroprotective, cardiant,
 CC antiinflammatory, antiparkinsonian and antiarthritic activity. The ptc-2

CC protein is a hedgehog receptor and is therefore capable of modulating
 CC hedgehog signalling, and so affect a number of hedgehog-mediated
 CC biological activities. The human patched-2 (ptc-2) protein can be used to
 CC screen for modulators, antagonists and agonists, which are likely to play
 CC an important role in the modulation of cellular proliferation and
 CC maintenance of, e.g. neuronal, testicular, osteogenic or chondrogenic
 CC tissues during disease states. Modulators of ptc-2 protein can be used
 CC for in vivo reformation of tissue; to improve grafting and morphology of
 CC transplanted tissue; for the treatment, prevention and/or reduction of
 CC the severity of neurological conditions deriving from: injury to the
 CC nervous system including traumatic injury, chemical injury, vascular injury
 CC and deficits (such as ischemia resulting from stroke), together with
 CC infectious/inflammatory and tumor-induced injury; ageing of the nervous
 CC system including Alzheimer's disease; chronic neurodegenerative diseases
 CC of the nervous system including Parkinson's disease, Huntington's chorea,
 CC amyotrophic lateral sclerosis, as well as spinocerebellar degenerations;
 CC and chronic immunological diseases of the nervous system including
 CC multiple sclerosis. ptc-2 therapeutics can also be used in the treatment
 CC of autonomic disorders of the peripheral nervous system, including
 CC disorders affecting the innervation of smooth muscle and endocrine
 CC tissue, e.g. to treat tachycardia or atrial cardiac arrhythmias.
 CC Antagonists of ptc-2 protein can be used to prevent differentiation of
 CC cells in culture, as well as for treatment of chronic pain syndromes.
 CC Agonists may be used to rescue neurons from lesion-induced death as well
 CC as neuron regeneration, in diseases such as CNS trauma infarction, (viral)
 CC infection, metabolic disease, nutritional deficiency, toxic agents, and
 CC so on. ptc-2 therapeutics may also be used for the repair of central and
 CC peripheral nerve damage, for repair and regeneration of non-neuronal
 CC tissue, e.g. skeletal and cartilage tissue, e.g. in the treatment of
 CC osteogenesis, arthritis, bone fractures, hereditary disease, as well as
 CC for generation of prosthetic cartilage devices, and to induce
 CC spermatogenesis and as fertility enhancers. This sequence represents the
 CC human ptc-2 protein described in the invention
 XX
 SQ Sequence 1203 AA;

Query Match 8.4%; Score 580; DB 2; Length 1203;
 Best Local Similarity 22.9%; Pred. No. 8.9e-44;

Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;

QY 306 RVAPARDKSKWDPKKGTSLSDKL--SPSTHTLLQGFPG-----WGTWVASHPLTILVLS 359
 Db 3 RSPPLRELPSPSTPPARTAAPOILAGSLKAPLWLRAYFQGLFLSGCGIQRHCGKVLPLG 62
 QY 360 VIPVVALAAGLVFTLTDPVELSAPNSQARSKAFHDHFG--PFRFTNOVILTAPNR 417
 Db 63 LIAFGALALGLRMALIEFNLEQLWVEVGSVQELHYHKEKLGEBEAAYTSQMLIQTRQE 122
 QY 418 SSRYDSSLGPKNPSGILDLLELLELLELRLHQLVQWSPQAQRNLSLDICY---AP 474
 Db 123 GEN-----LHQAALTASKVQVSLYKGSWDLNKICYKSGVP 166
 QY 475 LNPDN-----TSLYDCCINSLLQYFQNNRTLLLTATQTLMGQTSQVDWKD----- 520
 Db 167 LTENGMIERMTEKLPFCVILTPLDCFWGAK---LQGGSAVLPGRPDQMTNLDPEQLLE 223
 QY 521 -----HFLYCANAPLTPKDGDTALALSCMADYCA 548
 Db 224 ELGPFASLEGFRELDDKAQVQYVGRPCCLHPDDLHCPPSAPNHHSRQAPNVHAHLSGGC 283
 QY 549 PVVFPF-----LAIGGYKGDYSE---AEALIMTF-----SLNYPAGDPRLAQAKL 591
 Db 284 HGFSHKPMHQBELLGGWARDPQCELLRABALQSTFLLMSPRQLVEHFRGVDYQTHDIGW 343
 QY 592 WEEAFLEWRAPORMAGMFQVTFTAERSLB-DEINRTTAEDLPFATS-----Y-640
 Db 344 SEEQASTVLOAQWRFFVOLAOEALPENASQQTHAFSSTTLDDI-LHAFSEVSAARVWGY 402
 QY 641 IVIFLYISLALGSYSSWSRVVDSKATLGLGAVVLGAVMAAMGFPSVIRSRSLVILQ 700
 Db 403 LLMLAYACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGLCALGLITFNNAATQ 457
 QY 701 VVPFLVLSVGADNIFVLEVQLRRPRPCEPREVHIGRALGRVAPSMLLCSLSEATCFPL 760

Db 458 VLPFLALGIGVDDVFL--LAHAFTALPGTLOERMGECLQRTGTSVLTSTINMAAFM 515
QY 761 GALTMPAVRTFALTSLGLAVILDLQMSAFVALLSLDSKQEARLDVCC----- 812
Db 516 AALVPIPALRAFSQAALVVGCTFVAVMLVFPAFLSLDLRRHCORLDLCCFSSPSCAQ 575
QY 813 ---VKPQBL-----PPFGGEGE-- 827
Db 576 VIQILPQELGDTGTPVGTIAHLTATVQAFTHCEASSQHVVTILPPOAHLVPPSPDPLGSEL 635
QY 828 -----LLG-----FFQKAYAPFLHWTIGVLLFLA 855
Db 636 FSPGSTRDLGQEBEETRQKAACKSLPCARWNLAHFARYQAPALLOSHAKAIVLVLFGA 695
QY 856 LFGVLSYNSCHTSVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTILGYNFS--SEA 914
Db 696 LGLSLYGATLVQDGLALTDDVPRGTKEHAFLSAQLRYFSL--YEVALVTQGGFDYAHQR 754
QY 915 GMAAICSSAGCMNFSFTQKIQVATFEPFQSYLAIPAS-----SWDDFDIMLTSPSCC--- 967
Db 755 AL-----FDLHQRFSSL-----KAVLPPPATQAPRTWLHYRNWLQGIQAAPDQ 798
QY 968 -----RLYISGNKDKFCPSTVNSLNCILKNCMSITMGSVRSVQEHKYLFPWELND 1018
Db 799 DWASGRITRHSYRNG--SED-----GALAYKLLIQTGAQBPLO--FSQLTTRKLVD 846
QY 1019 RPNIKCPKGLAAYSVNLTSVGQVLASRFMAYHKP-----1055
Db 847 REGLIPPE--LFYMGILTVMVSSDPLGLAASQANFYPPPEWLHDKYDTTGENLRIPPAQP 904
QY 1056 -----LKNQSDYTEALRAARELAANI--TADLRKVPCTDPAPFVFPYTTINVPY 1102
Db 905 LEPAQPFLLRLGLQKTADFEATEGARACACAEAGQAGVHAYPSGSPF-----LFW 954
QY 1103 EQYLTILPGLMELSLC--LVPTFAVSCILGLDLSRGLNLSIVMILVDTVGFPMALWD 1160
Db 955 EQYLGIL--RRCFLAVAILLVCTFLCALLLNFWMAGLI--VLVLAAMTVELFGIMGFLG 1011
QY 1161 ISYNAVSLNVLAVAGMSVEFVSHITRFAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
Db 1012 IKLSAIPVVLVASGVGEFTVHVALGF--LTTQGSRLRAHALEHTAPVTDG--AIST 1069
QY 1221 LPGLVLGLAKAOLIQIFPRLMLLTLLGLLHGLVFLPVILSYVCPDVPNALALEOKBA 1280
Db 1070 LLGLLLAGSHFDFIVRYFPAALTVLTLGLLHGLVLLPVLSILGPP--PEVIQMYKES 1127
QY 1281 EBAVAV-----WVASCNHPRSVSTADNIYVN 1308
Db 1128 PEILSPAPOGGGLRWGASSLPQSFAVTTTSMVAIH 1165
RESULT 61
RAY92225
ID AAY92225 standard; protein; 1146 AA.
AC AAY92225;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human patched 2 (PTCH2) protein.
XX
KW Patched 2; PTCH2; tumour suppressor; familial melanoma CMM1;
KW familial adenomatous polyposis; hMomi; Michelin tire Baby Syndrome;
KW sonic hedgehog; SHH; development; carcinogenesis; cycostatic;
KW basal cell carcinoma; chromosome 1p32-35; gene therapy.
XX
OS Homo sapiens.
XX
FN WO200020037-A1.
XX
PD 13-APR-2000.
XX

PF 06-OCT-1999; 99WO-SE001784.
XX
PR 06-OCT-1998; 98SE-00003393.
XX
PA (KARO-) KAROLINSKA INNOVATIONS AB.
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Zaphiropoulos PG, Unden AB, Toftgard R, Rahnama F;
PI Hollingsworth RE;
XX
DR WPI; 2000-303645/26.
DR N-PSDB; AAA09081, AAA09084.
XX
PT Isolated human protein capable of participating in human patched
PT gene/sonic hedgehog pathway during embryonic development is used in
PT medicament for treatment of condition involving tumors such as basal cell
PT carcinoma.
XX
PS Claim 2; Page 41-44; 55pp; English.
XX
CC This protein is encoded by the novel human patched 2 gene (PTCH2), which
CC has been localised by radiation hybrid mapping to chromosome 1p32-35 with
CC DIS211 and W11404 as closest flanking markers and with an estimated
CC localisation 5.5 cR from DIS443. This region is often lost by LOH in
CC various different tumour types, such as neuroblastoma, melanoma and
CC breast and colon cancer. PTCH2 is a candidate gene for involvement in
CC gene in this region. It is also a candidate gene for involvement in
CC familial melanoma CMM1, modifier locus for familial adenomatous polyposis
CC hMomi and Michelin Tire Baby Syndrome. Splice variants of the PTCH2 gene
CC exist (see AAA09082-84). PTCH2 is capable of participating in the human
CC patched gene/sonic hedgehog (PTCH/SHH) pathway during embryonic
CC development and/or carcinogenesis. The isolated human protein is useful
CC as a medicament for the treatment of a condition involving tumours such
CC as BCC (basal cell carcinoma). The nucleic acid is useful in gene
CC therapy, and for use as a probe, primer or a diagnostic agent
XX
SQ Sequence 1146 AA;
Query Match 8.4%; Score 579; DB 3; Length 1146;
Best Local Similarity 23.3%; Pred. No. 1e-43;
Matches 276; Conservative 161; Mismatches 452; Indels 298; Gaps 40;
QY 306 RVAPARDKSMVDPKKGTSLSDKL--SPSTHTLQOPFQ--WGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSTPTAPAPQILAGSLKAPLWRAVFGLLFSLGCGIQRHCGKVLFLG 62
QY 360 VIPVVALAAGLVFTTLPDVELMSAPNSQARSEKAFHDQHF--PFRTNQVILTAPNR 417
Db 63 LLAFGALALGLRMATETNLSQLWVEGSRVSQELHYTEKLGEEAAYTSQMLIQARQE 122
QY 418 SSVRYDSLILGPKNPSGILDLDLLELLELQERLHLQVWSPQAQRNLSLODICY---AP 474
Db 123 GEN-----LHQAALTASKVQVSLYKSWDLNKKYKSGVP 166
QY 475 LNPDN-----TSLYDCCINSLQVQNNRTLLLTANTQLMGQTSQVMDKD----- 520
Db 167 LTENGMIERMIEKLPFCVILTDLDFEWGAK---LQGSAYLPGRPDQIQTWLDPEQLLE 223
QY 521 -----HFLYCANAPLTFKDGKTALALSCLMADYGA 548
Db 224 ELGPASLEGFRELIDKAQVQYVGRPCPLHPDDLHCPPSAPNHHSRQAPNVAHLSGGC 283
QY 549 PVPPPP-----LAIGYKGDYSE---AEALITWF-----SLNNYPAGDPRLLAQAKL 591
Db 284 HGFSHKFMHQBELIIGGMARDPQGBELLRAELAQSTFLMSPRQYEHFRGDIYQTHDIGW 343
QY 592 WEEAFLEEMRAFORRMAGMFQVTTFAERSLE--DEINRTTAEDLPFPATS-----Y 640
Db 344 SEEQASTVIOAWORRFVQLAQEALPENASQQIHAFSSITLDDI--LHAFSEVSAARVVG 402
QY 641 IVIFLYISIALGASYSSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFSSVILGRSSVILQ 700
Db 403 LLMLAVACTVM---LRWD--CAQSQSGVGLGVLVALAVASGLGICALLGITTFNAATTQ 457

Db 528 P-----FSLDVRFNEDETOQIASREFLIQAVNITDTHHEKEMVR----- 565
QY 1079 DLRKVPGETP--AFEVFPYTTITNVEQYLTPLEGFLMLSLCLVPTTFAVSCLLGLDLR 1136
Db 566 DLRQICKDPLNASIFHPYP---VFQDFELVRPVSQAMVIGAILMMIISFVIP-NIL 621
QY 1137 SGLNLLSIVMLVDTVGFMALWDISYNVSLINLVSAGMSVEFVSHITSPALSTKPT 1196
Db 622 CSLWAFVSISTELGVAGYMALWDVNLDSISMINIMCIGFSVDFTAHICYTMSKKRS 681
QY 1197 WLERAKEATINGSVAFFAGVATNLPGLTILVLGLAKAQLIQIFFFRNLNLLTLGLLGLV 1256
Db 682 PKARVREALHSLUGLPIIQSSSTIL-GIVALLAQSYIP-IVFFKXVFLVIFFGAMHGLF 739
QY 1257 FLPVILSYVGP 1267
Db 740 LLPVLLSLRGP 750
RESULT 63
ID AAW72971
AAW72971 standard; protein; 1311 AA.
XX AAW72971;
XX
DT 17-OCT-2003 (revised)
DT 26-JAN-1999 (first entry)
XX
DE DE
XX
KW Patched gene; ptc; diagnosis; treatment; developmental disorder; cancer;
KW healing; injured tissue; spina bifida; Wnt-1 oncogene; sperm production;
KW gene therapy.
XX
OS Junonia coenia.
XX
FH Key Location/Qualifiers
FT Misc-difference 348
FT /note= "unspecified"
FT Misc-difference 908
FT /note= "unspecified"
XX
FN US5837538-A.
XX
PD 17-NOV-1998.
XX
PF 06-OCT-1995; 95US-00540406.
XX
PR 07-OCT-1994; 94US-00319745.
XX
(STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Scott MP, Goodrich LV, Johnson RL;
XX
DR WPI; 1999-023461/02.
XX
XX Nucleic acid encoding vertebrate patched protein and related
PT transformants - used to express poly:peptide(s), useful for diagnosis and
PT treatment of developmental disorders or cancer, and in healing of injured
PT tissue.
XX
XX Example; Col 25-32; 38pp; English.
XX
XX The present invention describes vertebrate and invertebrate patched (ptc)
XX genes. Cells containing and expressing the ptc gene are used for the
CC recombinant production of the protein. These in turn are useful: (i) for
CC generating antibodies (Ab); and (ii) to screen for specific-binding
CC ligands (potential therapeutic agonists and antagonists). The ptc gene,
CC or its fragments, are used to isolate related sequences from other
CC mammals; to identify mutations (particularly those associated with
CC genetic diseases such as spina bifida and other developmental disorders);
CC to monitor expression levels in testis (to determine relationship with
CC sperm production) and to isolate 5'-non-coding sequences (used to study

CC embryonic development and to provide regulated expression of proteins).
CC The complete gene can be used in gene therapy, including expression of
CC antisense molecules, and to generate transgenic animals for studies of
CC embryonic development. Ab are used diagnostically to determine the ptc
CC protein on cell surfaces and as competitive inhibitors of signal
CC transduction through the ptc ligand. Cells that have been engineered to
CC express the ptc protein can be used to promote regrowth and healing of
CC damaged tissue (e.g. growth of new teeth) and regulation of the ptc
CC protein expression may be useful in cancer treatment (it may control the
CC Wnt-1 oncogene). The present sequence represents *Precis coenia*
CC (butterfly) patched amino acid sequence, from the present invention.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1311 AA;
Query Match 7.6%; Score 526.5; DB 2; Length 1311;
Best Local Similarity 22.2%; Pred. No. 1.1e-38;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;
QY 336 LLGQFFQGWGTWVASWPLTILVSVIPVVALAAGLVFTLTTPVELMSAPNSQARSEKA 395
Db 72 ILGCFLOGDAGKVL F--VAILVLTFCV-----GLKSAQIHTRVDQLWVQGGRLAEALK 124
QY 396 FHDQHFQFP-FRTNQVILTAPNRSYRYSLLLGPKNPFSGILDLDLLELLELLELRLHL 454
Db 125 YTAQALGEADSSHTQLVI---QTAKDPDVSLHP---GAL-----LEHL 162
QY 455 QVMSPEAQRNI-----SLQDICVAPLNPD-----NTSLYD---CCINSLLQYFQNN 497
Db 163 KVVHAATRVVHMVDIEWRKDKCYSPIPDFEGYHHIESITDNPVICAITPLDCFWEG 222
QY 498 RTLL-----
Db 223 SKLLGPDYPIVPHLKHKLQWHLNPLEVVEVKLKFQFPPLSTIEAYMKRAGITSAYMK 282
QY 516 ---VDWKDHFLYC-ANAPLTFKDGTFALALSCWADYGAPVFPF-----LAIGG----- 558
Db 283 KPCLDPTDP--HCPATAP-NKKSCHI PDVAAELSHGCGFAAAVYMHVPEQLIVGATRNS 339
QY 559 ----YKGDYSEAEALIMTFSLNYPAGDPRLAQAKLEEAFLSEMRAPFORMAGMFQVT 614
Db 340 TSALRKAEKLTQTVQLMGEREMYEYADHYKVHQIGWNOEKAAAVLDARQKFAAEVRKI 399
QY 615 FT-----AERSLEDEINRTTAEDLPIFATSVIVIFLYISIALGYSYSSWRVM 661
Db 400 TTSGSVSAYSPYFSTSTINDILGKFSVLSKNILGYMFMLIYVAVTL----IQW-RDP 455
QY 662 VDSKATLGLGVAVVLGAVMAAMGPFYSYLGRSSILVILQVVPFLVLSVGADNIFVLEY 721
Db 456 IRSQAGVGIAGVLLLSITVAAGLGFALLGIPFNASTQIVPFLALGLGVQDMFLTHTY 515
QY 722 QRLPRPGEPRVHIGRALGRVAPSMLLCSLSEACFFLGLGALTPMPAVRTEALTSLGLVI 781
Db 516 --VEQAGDVPREERTGLVLKKSGLSVLASCNVMAFLAAALLPIPAFRVFLQAAIILL 573
QY 782 LDFTLLQMSAFVALLSLDSKROBSRLDVCCKVKPQELPP----- 820
Db 574 FNLGSILLVFPAMISLDLRRSAAADLLCCLMPSPLPKKIPIERAKTRKNDKTHRIDT 633
QY 821 -----PGQEGEGL-----LLGFFQKAYAPFLHWHITRGVVLVLLFLALFVSLYSMC 865
Db 634 TRQPLDPDVSENVTCTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTYSWGTAT 693
QY 866 HHSVGLDQELALPKDSYLLDLYFLNLR---YFEVGAPVYFVTTLGNFSSAEAGMNAICSS 922
Db 694 KYVQGLDLTDIPENT---DEHEFSRQEKYFGF-----YN-----MYAVTQ- 732
QY 923 AGCWNFSF--TKIQVATEFPQSYLAIP-----ASSWVDDFIDWL----- 961
Db 733 ---GNFEYTNQKLLY--EYHDQ-FVRIINIKNDNGGLTKFWLSLFRDWLLDLQVAFDK 786
QY 962 -TPSSCCRLYISGFNKKDFCPSTVNSLNCIKMCMSTMGSVRPSVEQ----- 1007

Db 787 EVASGCI-----TQEWCKNASDE-GILAYKLMVQTGHVDNPIDKSLITAGHRLVDKD 838
 QY 1008 -----FHKYL-PWFLNDRPNKCPKGGI-----AAYSTSVNLTSDGOVLASRFM 1050
 Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPOQORWIHSPEDVHLEIKKSSPLIYTQLP 898
 QY 1051 AYHKPLKNSODYTEALRAARELANITADLRKVPGGTDPAPFVPPYITITNVFVQYLTILP 1110
 Db 899 FYLSGLSDTXSIXTKLRSVRDLCLKYE--KGLPN-----PPSGIPFLFMEQYLYLRT 949
 QY 1111 EGLFMLSCLVPTFAVSCLLGLDLSGLNLISIVMLIVDTVGFMAWDISYNAVSLIN 1170
 Db 950 SLLALACALAAVF-TAVWVLLNAAVAVLVTLALATLVQLLGVNALLGVLSAMPVAVL 1008
 QY 1171 LVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAVAMTNLPGLVGLA 1230
 Db 1009 LVLAIGRGVHTVHLCLGFTSICG---KRRASLALES-VLAPV---VHGALAAALA 1059
 QY 1231 KQAL-----IQIFFRNLITLGLLHGLVFLPVILSYVGP--DVNPALALEOKRAE 1281
 Db 1060 ASMLAASECGFVARFLRLLDIVFLGLIDGLLFFPPIVLSILGFAAEVRPI-----1110
 QY 1282 EAAVAVMVASCPNHRSRVST 1301
 Db 1111 -----EHPERLST 1118

RESULT 64

AAB67156
 ID AAB67156 standard; protein; 1311 AA.

XX AAB67156;

DT 12-APR-2001 (first entry)

XX Butterfly patched protein.

XX Butterfly; patched; PTC; segment polarity; limb patterning; development;
 KW hedgehog; antibody; human; mouse.

XX Junonia coenia.

XX US6172200-BI.

XX 09-JAN-2001.

XX 20-OCT-1997; 97US-00954668.

XX 07-OCT-1994; 94US-00319745.

XX 06-OCT-1995; 95US-00540406.

XX (STRD) UNIV LELAND S STANFORD.

XX Scott MP, Goodrich LV, Johnson RL;

XX WPI; 2001-136884/14.

XX N-PSDB; AAF32177.

PT Novel monoclonal antibody useful in diagnostic assays for detection of
 PT presence of protein on surface of cells specifically binds to naturally
 PT occurring patched protein, other than Drosophila patched protein.

XX Disclosure; Col 23-30; 39pp; English.

XX The present invention provides a monoclonal antibody which specifically
 CC binds to a patched protein (PTC) other than that from Drosophila. Also,
 CC given are the protein and coding sequences of patched from the beetle,
 CC mouse, butterfly, fruit fly, mosquito and human. Patched is a segment
 CC polarity gene involved in limb patterning. The sequences can be used to
 CC study development and to isolate the patched ligand, hedgehog. In
 CC addition, antibodies can be used to detect the PTC protein on cell
 CC surfaces or to inhibit the transduction of signal by the PTC ligand by
 CC competing for its binding site

XX SQ Sequence 1311 AA;
 Query Match 7.6%; Score 526.5; DB 4; Length 1311;
 Best Local Similarity 22.2%; Pred. No. 1.1e-38; Indels 307; Gaps 46;
 Matches 258; Conservative 175; Mismatches 420;
 QY 336 LLGQFPQGMGTWASWPLTILVLSVIPVVALAAGLVFTELTTDPVELWSAPNSQAKSEKA 395
 Db 72 ILGCFIQGDAGKVLFP--VAILVLSFVCV-----GLKSAQIHRTVDQLWVQEGRLAEALK 124
 QY 396 FHDQHPGPF-FRTNVILTPAPNRSYRYSLLIGPKNFGILDLDLLELLELELLELRLHL 454
 Db 125 YTAQALGEADSSHTQLVI---QTKADPDVSLIHP---GAL-----LEHL 162
 QY 455 QWSPAEQRNI-----SLQDICVAPLNPD-----NTSLYD-----CCINSILQVQNN 497
 Db 163 KVHAATRVTVHMYDIEWRLKOLCYSPSPDPFEGYHHISIIDNVIPCAIITPLDCFWEG 232
 QY 498 RTLL-----SLQDICVAPLNPD-----NTSLYD-----LLTANOTLMQOTSQ--- 515
 Db 223 SKLIGPDYPIYVPHLKHKLQWTHLNPLEVVEEVKKLKQFPLSTIEAYMKRAGITSAYMK 282
 QY 516 ----VWKDHLFYC-ANAPLTFKDGITALALSCWADYCAPVPPF-----LAIGG----- 558
 Db 283 KPCLDPTDP--HCPATAP-NKESGHIPDVAABELSHGCGYFAAAYMHWPEQLIVGGATRNS 339
 QY 559 ----YKGDYSEABALIMTFSLNNYPAGDPRLAQAQKLWEAFLEEMRAFORRMAGMFPQVT 614
 Db 340 TSALRRARXLQTVVQLMGEREMYWADHYKHVQIGNQOEKAAAVLDAMQKFAAEVRKI 399
 QY 615 FT-----AERSLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSSRSRVN 661
 Db 400 TTSGSVSSAYSFPFSTSTLNDILGKFEVSLKXIIILGTFWFLIYVAVTL---IQW-RDP 455
 QY 662 VDSKATLGLGVAVVLGAVMAAGPFSYLGIRSSLVILQVVPFLVLSVGADNIPFIVLEY 721
 Db 456 IRSQAGVGIAGVLLLSITVAAGLGFALLGIPFNASSTQIVPFLALGLGVQDMFLTHTY 515
 QY 722 QRLPRPGEPREVIHGRALGRVAPSMILCSLSAEACFPLGALTTPMNAVTFALTSLGLAVI 781
 Db 516 --VQAGDVPRBERTGLVLKKSGLSVLLASLQNVMAFLAAALLPIPAFRVFCLOAAILLL 573
 QY 782 LDFLQLMSAFVALLSDSKRQBASRLDVCCKVKPQELPP-----820
 Db 574 FNLGSLILVFPAMISLDLRRSARADLLCCLMPESPPLPKKIPERAKTKNDKTHRIDT 633
 QY 821 -----PGQEGE-----LLGFFQKAYAPFLHWHITRGVVLFLFALFGVLSYMC 865
 Db 634 TRQPLDPDVSENVTCCLSVSLTKWAKQYAPFIMRPAVKVTSMLALIAVILTSVMGAT 693
 QY 866 HISVGLDQELALPKDSYLLDYFLNR---YFVGVGAPVYFVTTLGVNFSSEAGMNAICSS 922
 Db 694 KVKOGLDLTDIVPENT---DEHEFLSRQEKYFGF-----YN-----MYAVTQ- 732
 QY 923 AGCNPNFSP--TOKIOVATEFPQSYLAIP-----ASSWVDDFIDLW-----961
 Db 733 ---GNFEPYTNQKLIY--EYHQDQ-FVRIENIINKDNGGLTKFWLSLFRDWLLDLQVAFDK 786
 QY 962 -TPSSCCRLYISGPNKDKFCPSTVNSLNCIKCMSITMGSVRPSVQ-----1007
 Db 787 EVASGCI-----TQEWCKNASDB-GILAYKLMVQTGHVDNPIDKSLITAGHRLVDKD 838
 QY 1008 -----FHKYL-PWFLNDRPNKCPKGGI-----AAYSTSVNLTSDGOVLASRFM 1050
 Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPOQORWIHSPEDVHLEIKKSSPLIYTQLP 898
 QY 1051 AYHKPLKNSODYTEALRAARELANITADLRKVPGGTDPAPFVPPYITITNVFVQYLTILP 1110
 Db 899 FYLSGLSDTXSIXTKLRSVRDLCLKYE--KGLPN-----PPSGIPFLFMEQYLYLRT 949
 QY 1111 EGLFMLSCLVPTFAVSCLLGLDLSGLNLISIVMLIVDTVGFMAWDISYNAVSLIN 1170

Db 950 SLLALACAAVF-IAVWLLNNAAVLTLALATLVLLQLLGVWALLGVLSAMPVL 1008
Qy 1171 LVSAVGSVEFVSHITSFALSTPTWLERAKETISMGSAVFAGVAMTNLPGILVLGLA 1230
Db 1009 LVLAIGRQVHTVHLCLGFTVTSIGC-----KRRASLALLES-VLAPV-----VHGALAAALA 1059
Qy 1231 KAOQL-----IQIFFRRLNLIITLLGLLHGLVLPVILSYVGP--DVPALALEQKRAE 1281
Db 1060 ASMLAASECGFVARLFLRLLDIVFLGLIDGLLFFPIVLSILGPAAEVRPI----- 1110
Qy 1282 EAVAAMVASCNHPSPVST 1301
Db 1111 -----EHERLST 1118

RESULT 65
AAG79573
ID AAG79573 standard; protein; 1311 AA.

AC AAG79573;

DT 29-AUG-2003 (revised)
DT 23-DEC-2002 (first entry)

XX Butterfly patched homolog.

XX Patched; ptc; basal cell nevus syndrome; BCNS; wound healing;
KW developmental disorder; spina bifida; craniofacial abnormality;
KW basal cell carcinoma; skin; brain tumour; transgenic; mouse;
KW proliferation; oncogenesis; signal transduction; cancer; aging.
XX Junonia coenia.

XX Key Location/Qualifiers
FH Misc-difference 348 /label= Unknown
FT Misc-difference 908 /label= Unknown
FT Misc-difference 908 /label= Unknown

XX US6429354-B1.

XX 06-AUG-2002.

XX 22-AUG-1997; 97US-00918658.

XX 07-OCT-1994; 94US-00319745.

XX 06-OCT-1995; 95US-00540406.

XX 31-MAY-1996; 96US-00656055.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX (REGC) UNIV CALIFORNIA.

XX Scott MP, Goodrich LV, Johnson RL, Epstein E;

XX WPI; 2002-722086/78.

XX N-PSDB; ABA00376.

XX Transgenic mouse, useful for cancer drug screening, comprises disruption
PT of allele of patched locus which predisposes the mouse to develop a
PT proliferative disorder.

XX Example; Col 47-54; 58pp; English.

XX This sequence shows the butterfly patched (ptc) protein homolog.
CC Mutations in the patched gene are characteristic in basal cell nevus
CC syndrome (BCNS), an inherited disorder with an increased risk of
CC developmental disorders such as spina bifida and craniofacial
CC abnormalities, basal cell carcinoma of the skin and brain tumours in
CC mammals. The human and mouse ptc sequences show 86% homology. The
CC transgenic mouse of the invention has a genome comprising a disruption of
CC at least one allele of a patched locus, where the disruption predisposes
CC the mouse to develop a proliferative disorder. The transgenic mouse is
CC useful for studying patched function and regulation, for e.g. a series of

CC small deletions and/or substitutions made in the patched gene, to
CC determine the role of different exons in oncogenesis, signal
CC transduction, etc; and for drug screening which is useful for treating
CC cancer or developmental abnormalities attributed to a defect in patched
CC function such as wound healing and aging. Note: The specification
CC indicates that this sequence is encoded by the cDNA given in ABA00376.
CC However no ORF which decoded to the corresponding protein could be found
CC in either orientation. (Updated on 29-AUG-2003 to standardise OS field)

XX SQ Sequence 1311 AA;

Query Match 7.6%; Score 526.5; DB 5; Length 1311;
Best Local Similarity 22.2%; Pred. No. 1.1e-38;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

Qy 336 LLGQFFQGWGTWVASWPLTILVSVTPVVALAGLVTELTTPVELWSPNQARSEKA 395
Db 72 ILGCFLLQGDAGKVLV--VAILVLTFCV-----GLKSAQIHTRVDQLMWOGGRLEALK 124
Qy 396 FHDQHFQFP-FRTNQVLTAPNRSYRYSLLGPKNFSGILDLDLLELLELQERLHL 454
Db 125 YTAQALGEADSTHQLVI---QTAKDPDVSLLHP-----GAL-----LEHL 162
Qy 455 QVMSPEAQRNI-----SLQDICYAPLNP-----NTSLYD-----CCINSLLQYQNN 497
Db 163 KVVHAATRVVMYDIEWRLKDLCSPSIPDFEGYHHIESIIDNVIPCAITPLDCFEW 222
Qy 498 RTLL-----LLTANQTLMGQTSQ--- 515
Db 223 SKLLGPDYPIYVPHLKHKLQWTHLNPLEVVEVKLFQPLPSTIEATMKRAGITSAVMK 282
Qy 516 ---VDMKDHFLYC--ANAPLTFKQGTALALSCMADYGAPVFPF-----LAIGG----- 558
Db 283 KPCLDPTDP--HCPATAP--NKSGHI PDVAELSHGCGYFAAYMHWPEQLIVGATENS 339
Qy 559 ---YKGDYSEAEALIMTFSNNYPAGDPRLAQAKLWEAEFLBEMAFORMAGMFQVT 614
Db 340 TSALRKARXLQTVVQLMGEREMYEYADHYKVHQIGWNQEKAAAVLDAMQKFAAEVRKI 399
Qy 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSWSRVM 661
Db 400 TTSQSVSSAYSFPFSTSTLNDILGKFSEVSKNIIILGYMFLYIVATL---IQW-RDP 455
Qy 662 VDSKATLGLGVAVVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLGFALLGIPFNASTQIVPFLALGLGVQDMFLITHTY 515
Qy 722 QRLPRRGPBPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVTFAITSLGLAVI 781
Db 516 --VEQAGDVPREERTGLVLKKSGLSVLASLNCVMAFLAALLPAPRFRVCLQAAILLL 573
Qy 782 LDFLLQMSAFVALLSLDSKQEAASRLDVCCCVKQPELPP----- 820
Db 574 FNLGSILVFPAMISLDLRRSAARADLLCCLMPESPPLPKKI PERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGE-----LLGFFQKAYAPFLLHMITRGVLLITRGVLLFLFALGVLSYMC 865
Db 634 TRQPLDPDVSENVTKTCCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTSVWGAT 693
Qy 866 HISVGLDQBELALPKDSYLLDYFLNLR---YFEVGAPYFVVTILGVNFSSEAGMNAICSS 922
Db 694 KVKDGLDLTDIVPENT---DEHEFLSRQEKYGF-----YN-----MYAVTQ- 732
Qy 923 AGCNFSPF--TKIOYATEFPQSYLAIP-----ASSWVDDFIDWL----- 961
Db 733 ---GNPEYPTNQKLLY--EYHQ--FVRIINIKNDNGGLTKFWLSLFRDWLLDLQVAFDK 786
Qy 962 -TPSSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVQ----- 1007
Db 787 EVASGCI-----TQEYCKNASDE--GILAYKLMVQTGHVDNPDIDKSLITAGHRLVDKD 838
Qy 1008 -----FKYLL--PWFINDRPNKCPKGGI-----AAVSTSVNLTSDDGQVLAASRFM 1050

Db 839 GIINPKAFYNYLSAWATNDALAYGASQGNLKPQQRWIIHSPEDVHLIHKSSPLIYTQLP 898
 QY 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEPVFPYITTNVFEQYLITLP 1110
 Db 899 FYLSGLSDTKSITKLIRSVRDCLCKYEA--KGLPN-----FPSGIPFLFWEQYLYLRT 949
 QY 1111 EGLFMLSCLVTPFAVSCLLGLDLSGLNLLSIWMILVDTVGFMAWDISYNVSLIN 1170
 Db 950 SULLALACALAAVF-IANVLLNAAVAVLTALATLVQLLGVNALLGVKLSAMPAYL 1008
 QY 1171 LVSAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVIGLA 1230
 Db 1009 LVLAIGRGVHTVHLCLGFVTSIGC-----KRRASLALES-VLAPV-----VHGALAAALA 1059
 QY 1231 KAQL-----IQIPFRMLNLLITLGLLHGLVFLFVLISYVGP--DVNPALALEQKRAE 1281
 Db 1060 ASMLAASECGFVARLFLRLLEDIVFLGLDGLLFFPPIVLISILGPAEVRPI-----1110
 QY 1282 EAVAAMVWASCPNHSRVST 1301
 Db 1111 -----EHPERLST 1118

RESULT 66

ABUG2149

ID ABUG2149 standard; protein; 1311 AA.

XX AC

XX ABUG2149;

XX DT 25-AUG-2003 (first entry)

XX DE Buckeye patched gene PTC product.

XX KW Buckeye; patched gene; PTC; hedgehog protein; gene therapy.

XX OS Junonia coenia.

XX FH Key

XX FT Misc-difference 348

XX FT /label= Unknown

XX FT Misc-difference 908

XX FT /label= Unknown

XX PN US2003032085-A1.

XX PD 13-FEB-2003.

XX PF 20-OCT-1997; 97US-00954701.

XX PR 07-OCT-1994; 94US-00319745.

XX PR 06-OCT-1995; 95US-00540406.

XX PA (SCOT/) SCOTT M P.

XX PA (GOOD/) GOODRICH L V.

XX PA (JOHN/) JOHNSON R L.

XX PI Scott MP, Goodrich LV, Johnson RL;

XX DR WPI; 2003-492065/46.

XX PT New DNA sequence other than present in a chromosome encoding patched gene

XX PT other than Drosophila patched gene, useful for preparing transgenic

XX PT laboratory animals and to knock out patched protein in embryonic stem

XX PT cells.

XX PS Disclosure; Page 8-10; 40pp; English.

XX CC The invention relates to a DNA sequence other than present in chromosome

XX CC encoding a patched (ptc) gene other than Drosophila patched gene. A cell

XX CC expressing the DNA sequence is useful for producing patched protein, by

XX CC growing the cell expressing the DNA sequence, where the patched protein

XX CC is expressed and isolating the patched protein free of other proteins.

XX CC The cell expressing the DNA sequence is also useful for screening

CC candidate compounds for binding affinity to the patched protein, by
 CC combining the candidate protein with the cell expressing the DNA sequence
 CC where the DNA sequence comprises the entire coding sequence under the
 CC transcriptional regulation of the transcriptional initiation region and a
 CC transcriptional termination region functional in the cell, expressing the
 CC patched protein in the cell and assaying for the binding of the candidate
 CC compound to the patched protein. The above cell is useful for screening
 CC candidate compounds for agonist activity with the patched protein. The
 CC DNA sequence is useful for producing all or portions of the patched
 CC protein, as probes for research, diagnosis, binding of hedgehog protein
 CC for its isolation and purification and in gene therapy. The DNA sequence
 CC is also useful as primers for investigating other species and for
 CC isolating genes from various mammalian sources of interest, particularly
 CC from humans or from domestic animals. The DNA sequence is further useful
 CC for preparing transgenic laboratory animals and to knock out the PTC
 CC protein in the embryonic stem cells, so as to produce hosts with single
 CC functional patched gene. The present sequence represents the amino acid
 CC sequence of the buckeye patched gene PTC product
 XX
 SQ Sequence 1311 AA;

Query Match 7.6%; Score 526.5; DB 7; Length 1311;

Best Local Similarity 22.2%; Pred. No. 1.1e-38;

Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 LLGQFFQCGMTWVASWPLTILVLSIPVVALAAGLVFTLTTPDVELMSAPNSQARSEKA 395
 Db 72 ILGCFLOQDAGKVLV--VAILVLSFCV-----GLKSAQIHTRVQDLMWQEGRELEAEIK 124
 QY 396 FHDHQFGFP-FRTNQVILTAPNRSRYRDSLLGPNFSGIULDLLLELLELQRLRLH 454
 Db 125 YTAQALGADSSHTQLVI---QTAKDPDVSLLHP---GAL-----LEHL 162
 QY 455 QWMSPEAQORNI-----SLQDICVAPLNPD-----NTSLVD---CCINSLIQVQNN 497
 Db 163 KVVHAARTVTVMYDIEWRLKDCYSPIDPEGVHHIESIIDNVIPCAITPLDQFWEG 222
 QY 498 RTLL-----LTTANQTLMGQTSQ-----515
 Db 223 SKLLGPDYPIYVPHLKHKLQWHLNPLFVVEVKLKFQPPPLSTIEAYMKRAGITSAYNK 282
 QY 516 ---VDWKDHFLYC-ANAPLTFKQGTALALSCMADYGAPVPPP-----LAIG-----558
 Db 283 KPCLDPTDP--HCPATAP-NKKSCHIPDVAEELSHGCGYGAAYAMHWPQLIVGGATRNS 339
 QY 559 ---YKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAELEEMRAFORWAGNQPVT 614
 Db 340 TSAURKARXLQTVQJMGEREYEWADHYKVHQIWNQOEKAAAVLDAMQKPAFAVRKI 399
 QY 615 FT-----AERSLEDEINRTTAEDLPIPATSYIVIFLYISALGSYSSWSRVN 661
 Db 400 TTSGSVSSAYSFPFSTSTLNDILGKPFSEVSLKNILGVMFMLIYAVTL---IQW-RDP 455
 QY 662 VDSKATILGGVAVVVLGAVMAAMGPFYSILGRSSILVILQVFPVLVSVGADNIFIPVLEY 721
 Db 456 IRSQAGVGIAGVLLLSITVAAGLGFPCALLGFPFNASTQIVPFLALGLGVQDMFLITHTY 515
 QY 722 QRLPRPGEPREHVHIGRALGRVAPSMALCSLSEACFFLGALTMPAVRTFALTSLAVI 781
 Db 516 ---VEQAGVDPREBRTGLVLKKSGLSVLASLCNVMAFLAAALLPIPAFRVFCQAAILLL 573
 QY 782 LDPLQLMSAFVALLSLDSKROEASRLDVCCCKVKPQELPP-----820
 Db 574 FNLSGILIVFPAMISLDLRRRSARADLLCCLMPESPLPKKKI PERAKTRKNDKTHRIDT 633
 QY 821 -----PQGGEG-----LIGFFQKAYAPFLHLLHITRGVILLFLFVGSLSYNC 865
 Db 634 TRQPLDPDVSENVVTKCTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLIALIVLTSSVWGAT 693
 QY 866 HISVGLDQELALPKDSYLLDYFLPLNR---YEEVGAPVYFVTLGYNFSSSEAGMAICSS 922
 Db 694 KYKQGLDLTDIVPENT---DEHFEFLSRQEKYFGF-----YN-----MYAVTQ- 732

XX The invention describes an assay for phenotyping the patched status of a
CC cell comprising detecting in a sample of mammalian cells the presence or
CC absence of a genetic lesion having aberrant modification or mutation of a
CC patched gene or mis-expression of the patched gene. The assay is useful
CC for diagnosing a genetic predisposition of an animal, e.g. basal cell
CC nevus syndrome, predisposition for developing tumour, i.e. carcinoma,
CC meningioma, medulloblastoma, or fibroma. A genetic construct encoding a patched
CC polypeptide is used to treat an animal having a disorder comprising loss
CC of function of a wild-type patched gene, such as cancer, and can enhance
CC patch function in e.g. wound healing and aging. This sequence is encoded
CC by the butterfly patched gene.

XX SQ Sequence 1311 AA;
Query Match 7.6%; Score 526.5; DB 8; Length 1311;
Best Local Similarity 22.2%; Pred. No. 1.1e-38;
Matches 259; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 LLGQFQGWGTWASWPLTILVLSVIPVVALAAGLVFTELTDPVELSAPNSQARSEKA 395
DB 72 ILGCFQGDAGKVPF--VAILVLSSTFCV-----GLKSAQIHTRVDQLWQEGGRLEALK 124
QY 396 FHDQHFQFPF--FRINQVILTAPNRSSYRYSDDLGPKNFSGILDLDLLELLELLELRLHL 454
DB 125 YTAQALGEADSTHQLVI---QTAKDPDVSLLHP---GAL-----LEHL 162
QY 455 QVMSPEAQRNI-----SLQDICYAPLNDP-----NTSLYD-----CCINSLLQYQNN 497
DB 163 KVVHAATRVTVHMYDIEWRLKDLCSYSPIDPFEGYHHIESIINDNVIPCAITLDCFWEG 222
QY 498 RTLL-----LLTANQITMGQTSQ--- 515
DB 223 SKLLGPDYPIYVPHLKHKLQWHLNPLEVVEVKLKFQPLSTIEAVMKRAGITSAYMK 282
QY 516 ---VDWKDHFLYC--ANAPLTFKQGTALALSCMADYDAPVFPF-----LAIG----- 558
DB 283 KPCLDPTDP--HCPATAP--NKSGHIDPVAEAELSHGCGYFAAAYMHWPEQLVGGATRNS 339
QY 559 ---YKQDYSEAEALIMTSLNNYPAGDPLAOKLWEAEFLERAFORMAGMFOVT 614
DB 340 TSALRKARXLQTVVQLMGEREMRYWADHYHQHJGWNQEKAAAVLDQWQKFAAEVRKI 399
QY 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISIALSGYSWSRVM 661
DB 400 TTSGSVSSAYSFPFSTSLNDILGKFSEVSLKNILGYMFMILYVAVTL---IQW-RDP 455
QY 662 VDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGAADNIFVLEY 721
DB 456 IRSQAGVGAGVLLSITVAAGLGFALLGIPFNASTSQIVPFLGALGVQDMFLTHTY 515
QY 722 QRLPRPCEPREVHIGRALGVAPSMILCSLSEACFFLGALTTPMPAVRTFALTSLGLV 781
DB 516 ---VEQAGDVPREERTGLVLKXSGSVLKLASLGNVAFLLAALLPAPFVFCVQLRAILL 573
QY 782 LDFLLQMSAFVALLSLDSKROEASRLDVCCVKKPQELPP----- 820
DB 574 FNLGSLILVFPAMISLDLRRSARADLLCCLMPESPFPKKKIPERAKTRKNDKTHRIDT 633
QY 821 -----PGQEGE-----LLGFFQKAYAPFLLHWTITRGVILLFLFGLSVLSYMC 865
DB 634 TRQPLDPDVSENVTKTCLSVLSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTSVWGAT 693
QY 866 HISVGLDQELALPKDSYLLDLYFLNRR---YFEVGAFFVYVTTGLYGFSSAGMNAICSS 922
DB 694 KYKQGLDLTDIVPENT---DEHFLSRQEKYFGF-----YN-----MYAVTQ- 732
QY 923 AGCNPF--TQKIQYATEFPEQSYLAIP-----ASSWDDFIWL----- 961
DB 733 ---GNFEYPTNQKLLY--BYHDQ-FVRIPIIKNDNGGLTKFMLSIFRDLWLLDLQVAFDK 786
QY 962 -TPSSCCRLYISGPNKDKFCPTSVNSLCLNCKMSITMGSVRPSVEQ----- 1007

Db 787 EVASGCI-----TOEYCKNASDE-GILAYKLMVQTHVNDPNIDKSLITAGHRLVDKD 838
 QY 1008 -----FKYL-PWFLNDRPNKCPKGG-----AAYSTVNLTSDQGVLASRFM 1050
 Db 839 GIINPKAFVNYLSAWATNDALAYGASQGNLKQPQORWIHSPEDVHLEIKKSSPLIYTQLP 898
 QY 1051 AYHKPLKNSQDTEALRAARELANITADLRKVGTDPAFVFPYTTITNVFEQVLTILP 1110
 Db 899 FYLSGLSDTXSKITLRSVRDCLKEYA--KGLPN-----FPGIPLFWEQVLYLRT 949
 QY 1111 EGLFWLSCLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGFMAWLDISYNAVSLIN 1170
 Db 950 SILLALACALAAVF-IAVNVLLNAAVAVLTLATLVQLGVMALLGVKLSAWPAVL 1008
 QY 1171 LVSAGVMSYFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLA 1230
 Db 1009 LVLAIGRGVHTVHLCLGPTVTSIGC-----KRRASLALES-VLAPV-----VHGALAALA 1059
 QY 1231 KAQL-----IQIFFRLNLLITLLGLLHGLVFLPVLISYVGP--DYNPALAUSQKRAE 1281
 Db 1060 ASMLAASECGFVARLFLRLLDIVFLGLTDLGLLFFPVLISILGPAAEVRPI-----1110
 QY 1282 EAVAAVMVASCPNHPSRVST 1301
 Db 1111 -----EHPERLST 1118

RESULT 70

ADN22811
 ID ADN22811 standard; protein; 1405 AA.

AC ADN22811;

DT 02-DEC-2004. (first entry)

DE Bacterial polypeptide #5464.

Recombinant DNA construct; transformed plant; improved plant property;
 cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 pathogen tolerance; pest tolerance; plant disease resistance;
 cell cycle pathway modification; plant growth regulator;
 homologous recombination; seed oil yield; protein yield; carbohydrate;
 nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 bacterial polypeptide.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 5464; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 1405 AA;

Query Match 7.4%; Score 514; DB 8; Length 1405;

Best Local Similarity 20.2%; Pred. No. 1.8e-37;

Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLSDKLSFSTHTLLGPFQCGWTV--ASWPLTILVLSVIVVLAAGLVFTELTTD 378

Db 100 RKGATGNRYALYSRSLIQKLLFALGNTVHRNAWSI-ILAVSMIFAVC-CYGLQYVHIETD 157

QY 379 PVELMSA-----PNSQARSEKAFHQHGPFF-RTN-----QVILTAPNRSS 419

Db 158 IVKLWVAOGGRDLBELNPLNIKEAMRNVTGDS--GPBLPENGGLGGYQVLIQTPEYEG 215

QY 420 YRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNI-----SLQDICY 472

Db 216 --QDALAAGP-----LLKHVEIMKHIAFNVSVHGVMDSLDICF 253

QY 473 APLNDP-----NTSLYD-----CC-INSLLQYFQNNRTU-----LLL 503

Db 254 KPAPPSVAADSAASLSDGVIDKIVPCITWITPIDCFWEGSKALGHPSPKSLGLPLGMLL 313

QY 504 TANQTLMGTSQVDMKD-----HFLYCANAPLTP-----532

Db 314 SS-----LSDGMIRWSDFDPIAVIDEIHRSFNLGSHYTFERAGVSHGMDRPCIDPLDP 369

QY 533 -----KDGITALALSCNADYGAPVPFPLAIGGYKGYSEAR 568

Db 370 ECPPMAKNYFDVCPHIDRVREIAKYGTELEBEKKKSGSPFDL---GRKKREAGDQ 426

QY 569 ALI-----MTFSLNYP-----AGDP 584

Db 427 KMIHPAQPADSIPTTEDAVPAQVPVSTAPITPTTTLSPPEARAAEKEKKQKARELKDYC 486

QY 572 -----KSLYNHKAENVDEGNERRTLH--PLASTSIADMLEBEFCQFNVTI 654

Db 487 KSYRKSAPFWLKKNDKWPVMSNNYFQNVDYAAEMTGGCGSPASVNLNWPEDMILGNP 546

QY 585 RLQAQ--AKL-----WEAFLEB-MRA 602

Db 547 RRAKGGKLSGADALQSVLPVLPADVFLRFKQKPGNRSMKTGLDMDAWNETAQVLOA 606

QY 603 FQRMAGMPQVTFTAERSL-----EDEINRTTAEDLPFATS-----YIV 642

Db 607 WOR-----NFT--KSLYNHKAENVDEGNERRTLH--PLASTSIADMLEBEFCQFNVTI 654

QY 643 IF-----LYISLALGSYSSWSRVWVDSKATLG--LGGVAVVLGAVMAAGFPYSLGIRSL 696

Db 655 ILGAYLMLAYAVITQARFDNCLPATESSMGLALAGVLVITFASVAGLGLATWFGIFBNA 714

Db 400 TTSGSVSSAYSYPSTSTLNDILGKFSVSLKNIILGYMFMLIYVAVTL-----IQW-RDP 455
QY 662 VDSKATGLGVAIVLGVMAAMGFYSYLGIRSSILVILQVFPFLVLSVGADNIFIVLEY 721
Db 456 IRSQAGVIGVLLSITVAAGLGFALLGIPNASSIQIVFPLALGLGVQDMFLTHY 515
QY 722 QRLPRPGPREVHIGRALGRVAPSMILCSLSEATCFPLGALTMPAVTTPALTSLGLAVI 781
Db 516 --VEGAGDVPREERTGLVKSGLSVLLASLGNVMAFLAAALLPAPRPFVCLQAAILL 573
QY 782 LQFLQMSAFVALLSDKQASRLDVCCKVPQELPP-----820
Db 574 FNLGSILVFPAMISLDLRRGAAPADLCCCLMPSPLEPKKIPERAKTRKNDKTHRIDT 633
QY 821 -----PGQEGEL-----LLGFFOKAYAPFLHWTIRGVLILLFALFGLSVLSYMC 865
Db 634 TRQPLDPDVSENVTKTCLSVSLTKWAKNQAFPIWRPAVKYTSMLALIAVILTSWGAT 693
QY 866 HISVGLDQELALPKDSYLLDYFLNLR-----YFEVGAPVYFVTTLGYNFSSSEAGMAICSS 922
Db 694 KVKDGLDLTDIVPENT---DEHEFLSRQEKYGF-----YN-----MVAVTQ- 732
QY 923 ACNNFSP--TKIQYATFPQSYLAIP-----ASSWDDFDWL-----961
Db 733 ---GNFEYPTNQKLLY--EYHQDQ-FYRIPNIIKNDNGGLTKFMLSFRDMLDLQVAFDK 786
QY 962 -TPSSCCRLYISGPNKDKPCPTVNSLNCMKMSITGWSVRPSVEQ-----1007
Db 787 EVASGCI-----TQEWCKNASDE-GILAYKLVQVGHVNDPNIDKSLITAGHRLVDKD 838
QY 1008 -----FKHYL-PMFLNDRPNIKCPKGL-----AAYSTVNLTSDDQVILASRPM 1050
Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPQORWIHSPEDVHLEIKKSSPLIVTQLP 898
QY 1051 AYHKELKNSQDTEALRAARELANITADLRKVGTDPAFVPPYTTITNVEQYLTILP 1110
Db 899 FYLSGLSDTDSIKTIRSVRDLCKLYEA--KGLPN-----FPSGIPFLFEQYL-YLR 948
QY 1111 EGLFMLSLCLVPTFAVSCLLGLDLSRGLNLLSVIMILVDTVGFMALWDISYNAVSLIN 1170
Db 949 TSLLALACALGAVFIAMVLLNAAVAVLVTLATVILQVLLGVKLSAMPPVL 1008
QY 1171 LVSAVGMSEVESHITRSPASTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLIA 1230
Db 1009 LVLAIGRGVHFTVHLCLGFPVTSIGC---KRRASLALBS-VLAPV---VHGALAALA 1059
QY 1231 KAQL-----IQIFFRLNLITLLGLLHGLVFLPVLISYVGP--DVNPALALSQKRAE 1281
Db 1060 ASMLAASEBGFVARLFLRLALLALVFLGLDGLLFFPVLISILGPAAEVRPI-----1110
QY 1282 EAVAAMVVASCPNHPRSVST 1301
Db 1111 -----EHPERLST 1118
RESULT 72
ABBS59092
ID ABBS59092 standard; protein; 1286 AA.
AC ABBS59092;
XX
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 4068.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL03195.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 4068; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1286 AA;
Query Match 6.9%; Score 474; DB 4; Length 1286;
Best Local Similarity 20.4%; Pred. No. 8.5e-34;
Matches 245; Conservative 177; Mismatches 457; Indels 322; Gaps 42;
QY 321 KGTSLSDKLSFSTHTLLGQFFQGMGTWVASWFLTILVLSVIPVVALAAGLVFELTTPDV 380
Db 43 KKGARGRTAIVLRSVFQSHLETLGSSVQKHAGKVLFAVILVLSFCVGLKSAQIHSKVH 102
QY 381 ELMSAPNQARSEKAFPHQHFQ-PFFRTNOVILTA---PNRSYRYDSLLGLPKNPSGIL 436
Db 103 QLWIOEGGRLEABLAYTKTIGEDSATHOLLIOTTHIDFNASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLHRHQVSPBAQRNLSLDICVAPLNPNTSLY-----DCC 486
Db 154 HLEVIVKATAVAKVHYDTE-W-----GLKDMCMFSTPSPFEGIYIEQILRHLPICS 204
QY 487 INSLIQYFQN-----NRTLILLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFWEGSQLGPSAVVIPGLNQLRLWTLNLPASVMQYKQKMSSEKISPDFTV 264
QY 522 FLYCANAPLTPKDGTA-LALSQM-----ADYGAPVFPPIAIG--GY 559
Db 265 EQYMKRAAI---GSGYMEKPCLNPLNPNCPDPAFNKNSTPDPVCA----ILSGCYGY 316
QY 560 KGDYSEAEALMTFSLNNYPAGDPLAQA-----589
Db 317 AAKMHMPEELIVGAKRN-RSGHLRKAQALQSVOQLMTEKEMYQMDQONKYVHHLGWTO 375
QY 590 -----KLWBEAF---LEEMRAFQRMAGMFQVTTFAERSLEDEINRTTARDLPFATS 639
Db 376 EKAAEVLNMQNRFSEVEQLLRKQSRATNVDIYVFPSSAALDDILAKFSHPSALSIVIG 435
QY 640 YIVIFLYTSLALGSYSWSRVNVDKATLGLGGVAVLGVANVAMGPFYSYLGIIRSLVL 699
Db 436 VAVTVLY---AFCTLLRW-RDPVRQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNAAST 491
QY 700 QVVPFLVLSVGADNIFIVLEYQRLPRPGPREVHIGRALGRVAPSMILCSLSEACFP 759
Db 492 QVVPFLALGLGVDFHIFMLTAAVYAESNRR--EQTKL-----ILKVGPSILFSACSTAGSPF 545
QY 760 LGALTMPAVRTFALTSGTLAVILDPLQMSAFVALLSLDSKQASRLDV-CCCV-----813

Db	546	AAAFIPVPALCVCLQAAIVMCSNLAALVFPAMISLDLRRRTAGRADIFCCCFPPWKE	605	PA	(GEMM/) GEMMILL R M.
Qy	814	KQOELPP-----PQOGEGL-----LLGF	831	XX	(DRAB/) DRABKIN H A.
Db	606	QPKVAPPVPLNNNGRGARHPKSCNNRVPLPAQNPLLEQORADIPGSSHSLASFSLATP	665	XX	Gemmili RM, Drabkin HA;
Qy	832	FQKAYAPELLHWITRGVVLLFLALFGVSLYSMCHISVGLDQELAPKDSVLLDYFLN	891	XX	WPI; 2002-712395/77.
Db	666	AFQHYTPFLMRSMVKFLTMGFLAALISLVASTRLODGLDIIDLVPKDSNEHKFLDAQT	725	XX	Novel Translocation in Renal cancer from Chromosome 8 genes, useful for
Qy	892	RYFEVGAPYFVITLGYNFSSEAGNNAICSSAGCNFSFTQIOVATEFPQSYLAIP--	949	XX	detection of tumors, comprises rearrangements in the t(3;8)(p14.2;q24.1)
Db	726	RLF--GFYSMAVTQG-----NFEYPTQQQLRDY-HDSFVRPHV	763	XX	chromosomal translocation which occurs in renal and thyroid carcinomas.
Qy	950	-----ASSWDDFIDWLTSPSSCCRLYISGPNKDKFCPSTVNSLNCXN-----	993	XX	Example 1; Fig 2C; 49pp; English.
Db	764	IKNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYRDGLTKCEWFPNASSD	811	XX	The invention relates to an isolated TRC8 (Translocation in Renal cancer
Qy	994	-----MSITWGSVRPSVEQ-----FKYL-PWFLNDRPNIKCPKG	1027	CC	from Chromosome 8) nucleic acid molecule, encoding a polypeptide
Db	812	AILAYKLIVQTGHVDNPDVKELVLTNRLVNSDGIINQRAFYNYLSAWATND-----	862	CC	comprising a sequence of 664 amino acids, fully defined in the
Qy	1028	GLAAYSTSNLTSQGVLASRPMAYHKPKNSQD-----YTEALRAARE	1071	CC	specificity and comprising a sequence located in the 5' flanking region
Db	863	-VFAYG-----ASQKLYPEPRQYHPQ--NEYDLKIPKSLPLVYAQMPFYLHGLTDSQ	914	CC	to the coding region of TRC8 and a sequence which occurs in certain
Qy	1072	LAANITADLRKVPGTDPFAVE--FPYITINVFYEQYLILPEGLFMLSCLCLVPTFANSCL	1129	CC	sporadic renal cell carcinomas. The methods are useful for detecting the
Db	915	I-KTLIGHIRDLVSKYEGFGLPNYPGIPFIWEQYMT-LRSSLAMILACVLLAALVLVS	972	CC	presence of the TRC8 gene in a biological sample, detecting alterations
Qy	1130	LLGLDLRSLGLNLLSIVMILVDTVGFMALWDISYNAVSLINLVSAVGSVFVSHITSF	1189	CC	to the gene, such as a 3;2 human chromosomal translocation, and fused DNA
Db	973	LLLLSWAAVVLVLSLASLAQIFGAMTLGKLSAIPAVILILSVGMWLCF--NVLISL	1030	CC	containing the fused site of TRC8/FHIT. A nucleic acid probe is useful
Qy	1190	AISTPTWLERAKETISGSAVFAGVAMTNLPGILVLGLAKAQIQLIFFRLNLLITLL	1249	CC	for detecting the 3;8 human chromosomal translocation, by contacting the
Db	1031	GPMTSVGNQRQVRQVLSQMQLGLPLVHGLMTSGVAVFMLSSTPFEVIRHFWLLLVLCV	1090	CC	nucleic acid probe with a biological sample to be tested, and determining
Qy	1250	GLLHGLVFLVFLVILSVYVGPDPVPALALEQKRAEAAVAVMVASCPNHPGRVSTADNIYNH	1309	CC	whether the nucleic acid probe specifically hybridises to the TRC8/FHIT or
Db	1091	GACNSLLVFPILLSMVGPEAE-LVPLE-----HPDRISTPSPLPVRS	1131	CC	FHIT/TRC8 fusion DNA. This sequence represents a TRC8 related protein of
Qy	1310	S 1310		XX	the invention
Db	1132	S 1132		XX	Sequence 1286 AA;
RESULT 73					Query Match
ABJ10929					6.9%; Score 474; DB 5; Length 1286;
ID	ABJ10929	standard; protein; 1286 AA.			Best Local Similarity 20.6%; Pred. No. 8.5e-34;
XX	AC	ABJ10929;			Matches 246; Conservative 176; Mismatches 465; Indels 306; Gaps 43;
XX	DT	12-DEC-2002 (first entry)			
XX	DE	TRC8 related protein Dm Ptc SEQ ID No 3.			
XX	KW	TRC8; Translocation in Renal cancer from Chromosome 8; fused DNA; 3,2;			
XX	KW	FHIT/TRC8 fusion DNA; sporadic renal cell carcinoma; TRC8/FHIT; TRC8FHIT;			
XX	KW	human chromosomal translocation.			
XX	OS	Drosophila melanogaster.			
XX	PN	US2002106656-A1.			
XX	PD	08-AUG-2002.			
XX	PF	02-JUL-2001; 2001US-00898533.			
XX	PR	12-MAR-1998; 98US-0077723P.			
XX	PR	12-MAR-1999; 99US-00268140.			
XX	XX				

```

QY 821 -----PQGBGL-----LLGPFQKAYAPF 839
Db 614' LPLNNNGRGARHPKSCNNRVPLPAQNPLEQADIPGSSHSLASFSLATPAFQHYTFP 673
QY 840 LLHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSVLLDYFLFLNRYPEVGAP 899
Db 674 LMRSWKFLTVMGFLAALSSLYASTRLQDGLDIDLVPKDSNEHKFLDAQTRLF--GPF 731
QY 900 VVFVTLTGNFNSSEAGMNAICSSAGCNFSFTQKIQIYATEFPEQSYLAIP----- 949
Db 732 SMVAVTQG-----NFEYPTQQLRDY-HDSFVRPHVKNKNGGL 771
QY 950 ASSWDDFDIMLTPSSCCRLYISGPNKDKPCSTVNSLCLNKC-----MS 995
Db 772 PDFWLLLFSEWL-----GNLQIFDEEYRDGRLTKECWFPNASSDAILAYKLI 819
QY 996 ITWGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGLAAYSTS 1035
Db 820 VOTGHVDNPNVDKELVLTNRLVNSDGLINQRAFTNLSAWATND-----VFAYG-- 867
QY 1036 VNLSDGQVLASRFMAYHKPLKNSOD-----YTEALRAARELANITAD 1079
Db 868 ---ASQKLYPFRQYFHP--NEYDLKTPKSLPLVYAQMPPYLHGLTDTSOI-KTILGH 921
QY 1080 LRKVPQTDPAFV--PPYTTNVFVEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRS 1137
Db 922 IRDLSVKYEGFGLPNYPGIPFWEQYMT-LRSSLAMILACVLLAALVLSLLLSVWA 980
QY 1138 GLNLLISIVMILVDTVGFALMDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTW 1197
Db 981 AVLIVLSLASLAQIFGANTLLGIKLSAIPAVILSVGMMLCF--NVLLISGFMTSVGN 1038
QY 1198 LERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQIQIFPFRNLMLTLGLLHGLVF 1257
Db 1039 RORRVQLSNQMSLGPLVHGMISGVAVFMLSIPFVIRHFCWLLLVLCVGCNSLIV 1098
QY 1258 LPVILSYGPDVNPALALQKRAEBAVAAMVASCNHPSPSVSTADNIYVNSH 1310
Db 1099 FBILLSWVGPEAB-LVPLE-----HPDRISTPSPLVRSS 1132

RESULT 74
AAR86304
ID AAR86304 standard; protein; 1299 AA.
AC AAR86304;
XX
XX 23-MAR-1996 (first entry)
XX
XX Drosophila patched protein.
XX
XX patched protein; hedgehog receptor protein; Drosophila; sonic hedgehog;
XX desert hedgehog; Indian hedgehog; drug screening; nervous system disease.
XX
XX Drosophila.
XX
XX Key Location/Qualifiers
XX Domain 93..426
XX /note= "Extracellular domain"
XX Domain 700..966
XX /note= "Extracellular domain"
XX
XX WO9518856-A1.
XX
XX 13-JUL-1995.
XX
XX 30-DEC-1994; 94WO-US014992.
XX
XX 30-DEC-1993; 93US-00176427.
XX 14-DEC-1994; 94US-00356060.
XX
XX (HARD ) HARVARD COLLEGE.
PA

```

(IMCR) IMPERIAL CANCER RES TECHNOLOGY.

Ingham PW, McMahon AP, Tabin CJ;

WPI: 1995-255060/33.

N-PSDB; AAQ91860.

Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful to treat degenerative nervous system disorder(s) and in gene therapy.

Disclosure; Page 170-76; 210pp; English.

The sequence represents a patched protein from *Drosophila*, which acts as a receptor for *Drosophila* hedgehog protein (AAR77337). The receptor or extracellular domain subfragments may be used in a drug screening assay, along with a vertebrate hedgehog protein, e.g. sonic hedgehog, desert hedgehog and Indian hedgehog proteins (AAR77338-R77341 and AAR77343-R77345). Compounds which inhibit or potentiate complex formation between the receptor and hedgehog protein may be used in therapy of neuromuscular, autonomic or central nervous system disorders

Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 2; Length 1299;

Best Local Similarity 20.4%; Pred. No. 6.7e-33;

Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

QY 321 KGTSLSDKLSFSTHTLLGQFQGMGTWASWFLTLVLVSVIPVWALAAGLVFTELTDPV 380

Db 43 KKGAGSRITAIYLRVSFQSHLETGSSVQKHAGKLVFVAILVLTFCVGLKSAQIHSKVH 102

QY 381 ELWSAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSYRYDSLLGLGPKNFGIL 436

Db 103 QLMIQEGGSLRAELAYTKTIGEDESATHQLLIQTHDPNASVLHPQALL-----A 153

QY 437 DLDLLELLELQERLHRHQVMSPEAQRNLSLODICVAPLNPNSTLY-----DCC 486

Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCMFSTPSPGIIYIEQILRHLPICS 204

QY 487 INSLIQYQN-----NRTLLALTAN-----QTLMGQTSQVDMKDH 521

Db 205 IITPLDCFWEGSQLLGPSAVVPIGLNQLRLTTLNPASVMQMKSEKISPDFTV 264

QY 522 FLYCANAPLTFKDGTAALSCM-----ADYGAPVPPFLAIG--GYK 560

Db 265 EQYMGRAIA---SGYMEKPCLNPNCPDPAPKNSTQPPDVG-----ILSGGCGYA 317

QY 561 GKDYSEABALIMTFSLLNNYPAGDPLQA-----589

Db 318 AKHMHPBELIVGGAKRN-RSGHLRKAQALQSVVQLMTEKEMYDQWQDNKYVHHLGWTQE 376

QY 590 -----KLWEERAF---LEEMRAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSY 640

Db 377 KAAEVLNAWQRFNSREVEQLRKQSIATNYDIYVFSSAALDDILAKFSHPALSIVIGV 436

QY 641 IVIFLYISLALGSYSWSRVVDSKATLGLGAVAVVLGAVMAAMGFFSYLGRSSILVIQ 700

Db 437 AVTVLY---APCTLLRW-RDPVRGQSSVGAVLLMCFSTAAGLGLSALLGIVFNAATQ 492

QY 701 VVPFLVLSVGADNIFIVLEYQRLPRRPGEPREVEHIGALGRVAPSMLLCSISEAICPPL 760

Db 493 VVPFLALGLGVDDHIFMLTAAYAESNRR--EQTKL-----ILKKVGPSILFSACSTAGSPFA 546

QY 761 GALTWPAPVTRPALTSLGLAVILDFFLOMSAFVALLSLDSKQREASRLDV-CCCV-----K 814

Db 547 AAFIEVPALKVFCLOAAIVMCSNTAAALLVFPAMISLDRRTTAGRADIFCCCFPMWKEQ 606

QY 815 POELPP-----PQGBGL-----LLGPF 832

Db 607 PKVAPVLPVPLNNGRGARHPKSCNNRVLPAPQNPLEQADIPGSSHSLASFSLATPA 666

QY 833 QKAYAPFLHMITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYFLFLNMR 892

Db 667 FQHTPTFLMRGWKFLTWGFLAALISSLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
Qy 893 YFEVGAPYFVTTLGYNFSSBAGNAICSSAGCANNFSGTQKIQYATFEPBQSYLAIP--- 949
Db 727 LF--GFYSWYAVTQG-----NFEYPTQQQLRDY-HDSFVRVPHVI 764
Qy 950-----ASSWDDFIDMLTFSSCCRLYISGPNKDKFCPSVNSLCLKNC----- 993
Db 765 KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYRDGLTKECWFNASSDA 812
Qy 994-----MSITMGSVRSVRSQ-----FHKYL-PWFLNDRPNIKCPRKG 1028
Db 813 ILAYKLIYQTHGVNDPNVDKELVTLNRLVNSDGIINORAFYNLSA-----ATNASSPEL 868
Qy 1029 LAA--YSTSVNLTSQGVLASRFMAYHKPLKNSQD-----YTBALRAAR 1070
Db 869 LRANCIRNRANGASQGLYPERQVYFHP--NEYDLKIPKSLPLVYAQMPPYLHGLTDT 926
Qy 1071 ELAANITADLRKVPGTDAFEV--FPYITVNFYEOYLTILPEGLFMLSCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDSVKYEGFGLPNYPSPGIPFIFWEQYMT-LRSSLAMTILACVLLAALV 984
Qy 1129 LILGLDRLSGLNLALSIVMILVDVTGFMALWDISYNAVSLINLVSAVGMSVEFVSHITRS 1188
Db 985 SULLSVWNAVILSVLASIAQIFGAMTLLGKLSAPAVILILSVGMCLCF--NVLIS 1042
Qy 1189 FAISTKPTWLREAKATISMSGSAVAGVAMTMLPGILVLGLAKAQIQLIFFRLMLLTL 1248
Db 1043 LGFMTSGVNRQRVQLSMQMSLGLVHGMLTSGVAVFMLSTSPFVFIRHFCWLLLVLC 1102
Qy 1249 LGLLGLVLPVLISVGVDPVNPALALEQKRAEEAAVAMVASCNPHSPRSTADNIYV 1308
Db 1103 VGACNSLLVFPILLSMVGPEAE-LVPLE-----HPRISTPSPLPVR 1143
Qy 1309 HS 1310
Db 1144 SS 1145

RESULT 75
ID AAB05389
XX AAB05389 standard; protein; 1299 AA.
XX AC AAB05389;
XX DT 12-SEP-2001 (first entry)
XX DE Human patched protein.
XX KW Human; hedgehog; morphogenic signal; neuron; patched protein;
KW embryonic patterning; cell culture; cell differentiation; ischaemia;
KW cell proliferative disorder; intracerebral grafting; Huntington's chorea;
KW neurological disorder; Alzheimer's disease; Parkinson's disease;
KW anyotrophic lateral sclerosis; ALS; multiple sclerosis.
XX OS Homo sapiens.
XX PN US6261786-B1.
XX PD 17-JUL-2001.
XX PF 02-JUL-1996; 96US-00674509.
XX PR 30-DEC-1993; 93US-00176427.
XX PR 14-DEC-1994; 94US-00356060.
XX PR 04-MAY-1995; 95US-00435093.
XX PR 05-JUN-1995; 95US-00460900.
XX PR 05-JUN-1995; 95US-00462386.
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
PA (HARD) HARVARD COLLEGE.
XX Marigo V, Tabin CJ, Ingham PW, McMahon AP;

XX WPI; 2001-440859/47.
DR N-PSDB; AAD10170.
XX
PT Screening compounds that potentiate or inhibit binding of hedgehog
PT polypeptide to naturally occurring patched receptor, comprises contacting
PT polypeptide with receptor and test compound, and detecting change in
PT binding.
XX
PS Disclosure; Col 193-202; 127pp; English.
XX
XX The present invention relates to assay for screening compounds that
CC potentiate or inhibit binding of hedgehog polypeptide to naturally
CC occurring patched receptor. The hedgehog proteins comprise morphogenic
CC signals produced by embryonic patterning centres, and are involved in the
CC formation and maintenance of ordered spatial arrangements of
CC differentiated tissues in vertebrates, both adult and embryonic. The
CC vertebrate tissues both in vitro and in vivo. The invention also relates
CC to a method for modulating growth, differentiation or survival of a
CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog
CC induction. Hedgehog agonists and antagonists can be used in cell culture
CC techniques to enhance survival and maintenance of neurons and various
CC determining organogenic pathways. The hedgehog gene is useful in
CC determining whether a patient is at the risk of disorder characterised by
CC unwanted cell proliferation or aberrant control of differentiation. The
CC hedgehog proteins or mimetics can be used to induce foetal neurons
CC especially neuronal stem cells in intracerebral grafting. The protein or
CC its mimetic can be used in the treatment of neurological conditions e.g.
CC injury to nervous system, ischaemia resulting from stroke, Alzheimer's
CC disease, Parkinson's disease, Huntington's chorea, anyotrophic lateral
CC sclerosis (ALS) and multiple sclerosis. The present sequence is human
CC patched protein related to the invention. Human patched protein is
CC provided in screening assays as a whole protein, or alternatively as a
CC fragment of the full length protein which binds to hedgehog proteins. The
CC patched protein can also be provided as a glycosylated protein
XX
SQ Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 4; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
Qy 321 KGTSLSDKLSFSTHTLLGQFFQGMGTWVASWPLTILVLISVIPVVALAGLVTELTDPV 380
Db 43 KKGARSGRTAIVLRVSFQSHLETLLGSSVQKAGKVLFAILLVSTFCVGLSAQIHSKVH 102
Qy 381 ELWSPAPSOARSEKAFHQHFG-PFRTNVOILTA--PNRSSRYVDSILLGPKPFGSL 436
Db 103 QLWIQEGGGLAEALAYTKTIGEDESATHQLLIQTHDFNASVLHPQALL-----A 153
Qy 437 DLDDLLELLELQERLRLHQLVWSPEAQRNLSQDICVAPLNPNTSLY-----DCC 486
Db 154 HLEUVLKATAVKVHLYDTE-W-----GLRDMCMSTPSPFEGIYIEQILRHLP 204
Qy 487 INSLILQYFQN-----NRTLLLLLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFWEGSQLLPESAVVIPGLNQLRLTTLNPASVMQMKQKMESEKISDFETV 264
Qy 522 FLYCANAPLTFKDGITALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db 265 EQYMKRAAJA---SGYMEKPCLNPLNPNCPDAPNKNSTQPPDVGA-----ILSGG 317
Qy 561 GKDYSEAEALIMTFSILNYPAGDPRIQA-----589
Db 318 AKHWHPEELIVGGAKEN-RSGHLRKAQALQSVVQLWTEKEMYDQDNVQVHHLGWTQE 376
Qy 590-----KLWEAP---LEEMRAFQRMAGNFQVTFATERSLEDEINRTTAEDLP 640
Db 377 KAAEVLNMQRNFSREVEQLRQSRATNIYDIYFSSAALDDILAKFHSPLSALSVIGV 436
Qy 641 IVTFILYISLALGYSYSSWSRVVDSKATLGLGVAVVGLGAVMAAMGFFSLGIRSSILVILQ 700

Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSGVAGVLLMCFSTAAAGLSALLGLGVFNASTQ 492
QY 701 VVPELVLSVADNIFVLEVYQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFL 760
Db 493 VVPELVLSVADNIFVLEVYQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFL 760
QY 761 GALTMPAPVTRTALISGLAVILDFLLQMSAFVALLSLSKQREASRLDV-CCCV-----K 814
Db 547 AAPIPVPALKVFCLOAAVIMCSNLAALVFPAMISLDRRTAGRADIFCCCFPVWKQ 606
QY 815 PQELPP-----PQGBGL-----LLGPF 832
Db 607 PKVAPVPLPNNNGRGARHPKSCNNRVALPAQNPLEQRADIPQSSHSLSASFSLATPA 666
QY 833 OKAYAPFLHWTGRGVVLLFLALFGVLSYMSCHISVGLDQBLALPKDSYLLDYFLFLNR 892
Db 667 FQHYTPFLMRSVKFLVTMGFLAALLISSLYASTRIQDGLDIIDLVPKDSNEHKLDAQTR 726
QY 893 YFEVGAPVYVTVTLGNFSSBAGMAICSSAGCNFSPFKIQYATEPPEQSYLAIP--- 949
Db 727 LF--GFYSMYAVTQG-----NFEYPTQQQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWDDFDIMLTSSCCRLXISGPNKDKFCFSTVNSLNLKNC----- 993
Db 765 KNDNGLPDFWLLFSEWL-----GNLQKIFDEEYRDGLTKECWFPNASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----PHKYL-PWFLNDRNRIKPKGG 1028
Db 813 ILAYKLIQVQGHVDPVKELVITNRLVNSDGIINQAFYNYLSAW-----ATNASSPTL 868
QY 1029 LAA--YSTSVNLTSGQVLASRFMAYHKPLNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQKLYPEPRQVPHQ--NEYDLKIPKSLPVYAQMPFYLHGLTDS 926
QY 1071 ELAANITADLRKVGTDPAFEV--PYTITNVFYQYLTILPEGLMFLSLCLVPTFAVSC 1128
Db 927 QI--KTLIGHIRDSLVKVEGFGNTPSPGIPFIFWESQYMT-LRSSLAMILACVLLAALVLV 984
QY 1129 LLLGLDLSGLLNLASIVMLVDTVGFMAWDISYNAVSLINLISAVGMSVFEVSHITRS 1188
Db 985 SLLLSVAAVILSVLASLAQIFGAMTLLGLKLSAPAVILISVGNMLCF--NVLS 1042
QY 1189 PAISTKPTWLERAKATISMGSAFVAGVAMTNLPGILVIGLAKAQLIQIFPFLMLLITL 1248
Db 1043 LGFMTSVGNRRQVQLSMQSLGPLVHGMLTSGVAVFMLSTSPFBEVIRHFCWLLLVLC 1102
QY 1249 LGLLHGLVPLVILSVGPDVNPALALEQKRAEEAAVAVMVASCPNHPRSVSTADNIYN 1308
Db 1103 VGACNSLLVFPILLSWVGPEAB-LVPLE-----HPDRISTPSPLPVR 1143
QY 1309 HS 1310
Db 1144 SS 1145
RESULT 76
ID AAB85751
XX AAB85751 standard; protein; 1299 AA.
AC AAB85751;
CC AAB85751;
DT 29-OCT-2001 (first entry)
XX Hedgehog (hh) polypeptide related sequence Id No. 48.
XX Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh; Dhh;
KW desert hedgehog; cell differentiation.
XX Unidentified.
XX US6271363-B1.
XX 07-AUG-2001.

XX 20-OCT-1997; 97US-00954698.
XX 30-DEC-1993; 93US-00176427.
PR 14-DEC-1994; 94US-00356060.
PR 04-MAY-1995; 95US-00435093.
PR 05-JUN-1995; 95US-00462386.
XX (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX Ingham PW, McMahon AP, Tabin CJ;
XX WPI; 2001-456723/49.
DR N-PSDB; AAH76131.
XX Novel nucleic acid encoding a hedgehog polypeptide, used to produce the
PT polypeptide, which is used to promote proliferation, survival, and/or
PT differentiation of neuronal and mesodermal tissue.
XX Disclosure; Col 181-188; 118pp; English.
XX The invention relates to nucleic acids encoding hedgehog proteins
CC selected from sonic hedgehog (Shh), indian hedgehog (Ihh), desert
CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the
CC formation of ordered spatial arrangements of differentiated tissue in
CC vertebrates. The nucleic acid sequences are useful for producing hedgehog
CC proteins, used for promoting differentiation of, or survival of
CC differentiated, neuronal cells, and for promoting proliferation, survival
CC or differentiation of mesenchymal, endodermal or ectodermal tissue,
CC particularly chondrocytes, or testicular germ line cells. The present
CC sequence represents a hedgehog polypeptide related sequence
XX Sequence 1299 AA;
SQ
Query Match 6.7%; Score 464.5; DB 4; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSPTHTLLGQFGQMGVWASWPLTILVLSVIPVVALAAGLVFTELTDV 380
Db 43 KKGAGSTAIYLRVSFQSHLETGLSSVQKHAGKLVFVAILVLSFTCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSYRYSLLGLPKNPSGIL 436
Db 103 QLMIQEGGGLAEALAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELLEQLRHLQVMSPEAQRNLSQDICVAPLNPDNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHLDYDE-W-----GLRDMCMFSTPSFEGIYIYIEQILRHLIPCS 204
QY 487 INSLLOYFQN-----NRTLALLTAN-----QTLMGQTSQVMDKDH 521
Db 205 IITPLDPCFWEGSQLGSPESAVVPIGLNORLLMTLLNPASVMQYMKQMSSEKISDFETV 264
QY 522 FLYCANAPLTKDGTALALSCM-----ADYGAPVPPFLAIG--GYK 560
Db 265 EGYMKRAAIA---SGYMEKPCLNPLNPNCPDAPNKNSTOPEVGA-----ILSGCYGTA 317
QY 561 GROYSEAEALIMTFSLNYPAGDPRLAQA----- 589
Db 318 AKHMHWPBELIVGGAKRN-RSGHLRKAQALQSVQVQLMTEKEMYDQMDQNVKVVHLGWTQE 376
QY 590 -----KLWEAP---LEEMRAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSY 640
Db 377 KAAEVNNAWQRNFSREVQQLRKQRIATNYDIYVFSSAALDDILAKFSHPALSIVIGV 436
QY 641 IVIFLYISIALGSSYSWSRVMVDSKATLGLGVAVVVLGAVMAAMGFFSVGLRSLSVLIQ 700
Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSGVAGVLLMCFSTAAAGLSALLGLGVFNASTQ 492
QY 701 VVPELVLSVADNIFVLEVYQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFL 760

493	QY	VVPFLALGLVDHIFMLTAAVAESNRR--EQTKL-----ILKKVGPSILFSACSTAGSFFA	546
761	QY	GALTMPAVRTFALTSGVLVDLFLOWSAFVALLSDSKRQEARLDV-CCCV-----K	814
547	Db	AAFTVPALKVFCLOAAIVMCSNLAALLVFPAMISLDLRRTAGRADIFCCCFPVWKEQ	606
815	QY	POELPP-----PQQEGIL-----LLGFF	832
607	Db	PKVAPPVLPNNNGRGARHKPSKNNNRVALPAQNPLEQRADIPGSSHSLSASFSLATFA	666
833	QY	QKAYAPFLLHWITRGVLLLFALPGVSLYSMSCHISVGLDQELALPKDSYLLDYFLFLNR	892
667	Db	FQHYTFPLMRWVKFLTWGFLWLAALISSLYASTRAILODGLDIIDLVPKDSNEHKFLDAQTR	726
893	QY	YFEVGAPVYFVTTLTGYNFSSEAGNMAICSSAGCANNFSFTQKIQVATEPPEQSYLAIP	949
727	Db	LF--GFYSMYAVTQG-----NFEYPTQQQLRDY-HDSFVRVPHVI	764
950	QY	-----ASSWDDFIDWLTPSSCCRLYISGNPKDKFCPSTVNSLNCIKNC-----	993
765	Db	KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYRDGRLTKECWFNASSDA	812
994	QY	-----MSITWGSVRPSVEQ-----FHKYL-PWFLNDRPNKCPKGG	1028
813	Db	ILAYKLIIVQTGHVDNPVDKELVLTWRLVNSDGIINQRAFYNYLSAW-----ATNASSPTEL	868
1029	QY	LAA--YSTSVNLTSDGOVLASRFMAYHKPLKNSQD-----YTEALRAAR	1070
869	Db	LRANCIRNPANGASQGLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQMPPFYHLGLTDTS	926
1071	QY	ELAAINITADLRKVPGTDPAFEV--FPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSC	1128
927	Db	QI--KTLIGHIRDLSVKYEGFGLNPFYPSGIPFIWFQYMT-LRSSLAMILACVLLAALVLV	984
1129	QY	LILGLDRLSGLANLISIVMLVDTVGFMAWDISYNAVSLINLVSAGVSVFVSHITRS	1188
985	Db	SLLLSVMAAVILVSLASLAQIFGAMTLLIGIKLSAIPAVILILSVGMMLCF--NVLIS	1042
1189	QY	FAISTKPTWLERAKEATISMSAVFAGVAMTNLPGILVLGLAKAQIQLIFFRLNLLITL	1248
1043	Db	LGPMTSVGNRRQRRVQLSMQSLGPLVHGMLTSGVAVFMLSTSPFEFVIRHFCWLLLVLC	1102
1249	QY	LGLLHGLVFLPVLLSVYGVDPNPALALEQKRAEEAAVAVMVASCPNHPSRVSTADNIYN	1308
1103	Db	VGACNSLLVPILLSMVGPEAE-LVPLE-----HPDIRSTPSPLPVR	1143
1309	QY	HS	1310
1144	Db	SS	1145
RESULT 77			
AAB31232			
ID	AAB31232 standard; protein; 1299 AA.		
XX	AAB31232;		
AC	AAB31232;		
XX	20-APR-2001 (first entry)		
XX	A consensus hedgehog-related protein.		
XX	Hedgehog related-protein; sonic hedgehog protein; Shh; ischemia; stroke;		
KW	desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;		
KW	neurological condition; nervous system injury; tumour-induced injury;		
KW	aging; Alzheimer's disease; chronic neurodegenerative disease;		
KW	Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;		
KW	spinocerebellar degeneration; chronic immunological disease;		
KW	multiple sclerosis.		
XX	Synthetic.		
OS	Unidentified.		
XX	US6165747-A.		

XX	PD	26-DEC-2000.	
XX	XX	05-JUN-1995;	95US-00460900.
XX	PR	30-DEC-1993;	93US-00176427.
XX	PR	14-DEC-1994;	94US-00356060.
XX	PR	04-MAY-1995;	95US-00435093.
XX	PA	(HARD) HARVARD COLLEGE.	
XX	PA	(IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.	
XX	PI	Ingham PW, McMahon AP, Tabin CJ, Marti-Gorostiza E, Bumcrot DA;	
XX	DR	WPI; 2001-079847/09.	
XX	DR	N-PSDB; AAC87096.	
XX	PT	Polynucleotides encoding hedgehog proteins, useful for treating diseases of nervous system such as Alzheimer's disease, Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple sclerosis.	
XX	PS	Disclosure; Col 183-192; 119pp; English.	
XX	CC	The present sequence represents consensus a hedgehog related-protein. The specification describes a sonic hedgehog protein (Shh), a desert hedgehog protein (Dhh), and an indian hedgehog protein (Ihh). The hedgehog polynucleotides are useful in diagnostic, in antisense therapy and in therapeutic assays for detecting and treating disorders involving, e.g., aberrant expression of vertebrate hedgehog homologue. Hedgehog polypeptides are useful therapeutically to enhance survival of neurons and other neuron cells and in treating neurological conditions deriving from acute, subacute, or chronic injury to the nervous system, including traumatic injury, chemical injury, vascular injury and deficits (such as the ischemia resulting from stroke), together with infectious/inflammatory and induced-induced injury, aging of the nervous system including Alzheimer's disease, chronic neurodegenerative diseases of the nervous system, including Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis, spinocerebellar degenerations, and chronic immunological diseases of the nervous system or affecting the nervous system, including multiple sclerosis	
XX	SQ	Sequence 1299 AA;	
Query Match			
Best Local Similarity 6.7%; Score 464.5; DB 4; Length 1299;			
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;			
QY	321	KGTSLSDKLSFTHTLQPFQCGHGTWVASWPLTILVLSVIPVVALAAGLVTELTDPV	380
Db	43	KGKAGSRTAIYLRSVFQSHLETGLSSVQKHAGKVLFAILVLSTFCVLGKSAQIHSKVH	102
QY	381	ELWSAPNSQARSEKAFHQHFG-PFTRTNQVILTA---PNRSSRYDSDLLGPKNPSGIL	436
Db	103	QLWIOEGGGLAEALAYTOKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A	153
QY	437	DLDDLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPNTSLY-----DCC	486
Db	154	HLEVLKATAVKVHLYDTE-W-----GLRDMCNMPSFSGFIYIEQILRLHIPCS	204
QY	487	INSLLOYFON-----NRTLLLLLTAN-----QTLMGQTSQVDWKDH	521
Db	205	LIITPLDCFWEGSQLLPESAVVIPGLNORLLTTLNPASVMQMKQMSSEKISPDFETV	264
QY	522	FLYCANAPITFDGKTALALSCM-----ADYGAPVPFFLAIG--GYK	560
Db	265	EQMKRAIAA--SGYMEKPCPLNPCLNPNCPTAPNKNSTQPPDVGA---ILSGGCGYA	317
QY	561	GKDYSEAEALIMTFSLNNYPAGDPRLAQ-----	589
Db	318	AKHMHWPPEELIVGGAKRN-RSGHLRKAQALQSIVVQLMTEKEMYDQMDNKKVHHLGWTOE	376
QY	590	-----KLWEEAF---LEEMRAFORMAGMCFQVTFABERSLDEINRTTAEDLPIFATSY	640

Db 377 KAAEVLNANQRFSEVEQLLRKQSRATNYDIYVSSAALDDILAKFHPHSALSIIVGV 436
 QY 641 IVIFLYISIALGSSYSWSRVMDSKATLGLGGVAVVVLGAVMAAMGPFYSYLGRSSIVLIQ 700
 Db 437 AVTVLY--AFCTLLRW-RDPVGGSSVGVAGVLLMCFSTAGLGLSALLGIVFNAASQ 492
 QY 701 VVPFLVLSGADNIFIVLEYQRLPRPQBPREVHIGRALGRVAPSMMLCSISEALCPFL 760
 Db 493 VVPFLALGLGVDFHFMFTAAYAESNR--EOTKL-----ILKKVGPISILFSACSTAGSFFA 546
 QY 761 GALTMPAVRTALTSGLAVILDELQMSAFVALLSLDSKROEASLDV-CCCV-----K 814
 Db 547 AAFIPVPAKLVFCLQAAIWMCSNLAAALVFPAMISLDDRRRAGRADIFCCCFPVWKEQ 606
 QY 815 PQELPP-----PQGGEGE-----PQGGEGE-----LLGFF 832
 Db 607 PKVAPPVLPLNNNGRGARHPKSCNNRVALPAQNPFLLEQRADIPGSSHSLASFSLATPA 666
 QY 833 QKAYAPFLHWTIRGVVLLFLALFCVLSYMSCHISVGLDQELALPKDSYLYLDYFLFLNR 892
 Db 667 FQHYTPFLMRSVKFLTVMGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
 QY 893 YFEVGAPVVTTLGYNFSSSEAGMNAICSSAGCNPFSFTQKIQYATEPPEQSYLAIP--- 949
 Db 727 LF--GFYSIAVYQG-----NFEYPTQQLRDY-HDSFVRVPHVI 764
 QY 950 -----ASSWDDFDIMLTPSSCCRLYISGPNKDKFCPTVNSLNLKNC----- 993
 Db 765 KNDNGGLPDFWLLFSEWL-----GNLQKIPDEBYRDGRLTKECFPPNASSDA 812
 QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRNIPKPKGG 1028
 Db 813 ILAYKLIVQTHVDNPDVKELVLTNRLVNSDGIINQAFNYLSAW-----ATNASSPTL 868
 QY 1029 LAA--YSTSVNLTSQVLAISRFMAVHKPKNSQD-----YTEALRAAR 1070
 Db 869 LRANCIRNRANGASQGLYPEPRQYHPQ--NEYDLKIPKSLPLVYAQMFPYHLGLTDS 926
 QY 1071 ELAANITADLRKVPGTDPAPFV--FPYITNVFYEQYLTPLEGFLMLSCLVPTPAVASC 1128
 Db 927 QI-KTLIGHIRDLVKYEGFGLPNYPSGIPFIFWEQYMT-LRSSLAMILACVLLAALVLV 984
 QY 1129 LLLGLDLRGLNLLSIVMLVDTVGFMAWDISYNAVSLINLVSAVGMSVRFVSHITRS 1188
 Db 985 SLLLSVMAAVLVLSLASLAQIFGAMTLGLKLSAIPAVILLISVGNMMLCF--NVLIS 1042
 QY 1189 PAISIKPTWLERAKEATISGSAVAGVAMTNLPGILVLGLAKAQILQIFPRALNLTIL 1248
 Db 1043 LGFMTSVGNRRVQLSMQMSGLPLVHGLMTSGVAFVMLSTSPFEFVIRHFCWLLLVLC 1102
 QY 1249 LGLLHGLVLPVILSYVGPDPVPNALALEQKRAEEAAVAVWASCPNHPRSVSTADNIYN 1308
 Db 1103 VGACNSLLVFPILLSVWGFEAR-LVPLE-----HPRISTPSPLPVR 1143
 QY 1309 HS 1310
 Db 1144 SS 1145

RESULT 78

ABB79150
 ID ABB79150 standard; protein; 1299 AA.

AC ABB79150;

XX 06-AUG-2002 (first entry)

XX Patched protein sequence SEQ ID NO:42.

XX Sonic hedgehog; Shh; desert hedgehog; Dhh; Indian hedgehog; Ihh;
 KW antiparkinsonian; antiarrhythmic; neuroprotective; anticonvulsant;
 KW cytostatic; neurotropic; spermatogenesis; peripheral nervous system;
 KW central nervous system; Alzheimer's disease; Parkinson's disease;

KW Huntington's disease; arrhythmia; nerve degeneration; multiple sclerosis;
 immunological disorder; neoplastic; hyperplastic.

XX Unidentified.

XX US6384192-B1.

XX 07-MAY-2002.

XX 20-OCT-1997; 97US-00957874.

XX 30-DEC-1993; 93US-00176427.

XX 14-DEC-1994; 94US-00356060.

XX 04-MAY-1995; 95US-00435093.

XX 05-JUN-1995; 95US-00462386.

XX (HARD) HARVARD COLLEGE.

XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

XX Ingham PW, McMahon AP, Tabin CJ;

XX WPI: 2002-442817/47.

XX N-PSDB; ABN87568.

XX New vertebrate hedgehog-related proteins, useful e.g. for promoting
 differentiation, survival and proliferation of cells, e.g. for treating
 neurodegeneration.

XX Disclosure, Col 171-180; 116pp; English.

XX The present invention describes an isolated and/or recombinant
 polypeptide (I) comprising a hedgehog (hh) amino acid (aa) sequence
 encoded by a nucleic acid (II) that hybridizes under stringent conditions
 to 1 of 6 sequences (see ABN87544, and ABN87546 to ABN87550). (I) binds
 to a natural patched receptor. Specifically claimed example of (I) are
 given in ABN79132 and ABN79134 to ABN79138. (I) has antiparkinsonian,
 neurotropic, neuroprotective, anticonvulsant, antiarrhythmic and cytostatic
 activities. (I) induces the expression of the BMP-2 and -4 genes, and of
 the Hoxd gene. (I) can be used: (i) to promote differentiation of
 neuronal cells and survival of the differentiated cells, specifically
 dopaminergic or motor neurons, proliferation of chondrocytes, and
 proliferation, differentiation and/or survival of mesodermal or
 ectodermal cells, either in cell cultures (particularly for preparation
 of transplants) or therapeutically; (ii) for detecting loss of response,
 in tissues or, to hh proteins; (iii) in drug screening (to identify
 (ant)agonists, useful e.g. for inhibition of spermatogenesis); and (iv)
 for isolation of cognate receptors. (I) may be used therapeutically to
 treat e.g. injuries/defects in the central or peripheral nervous systems,
 including Alzheimer's, Parkinson's and Huntington's diseases, or
 arrhythmias caused by nerve degeneration; immunologic disorders of the
 nervous system, e.g. multiple sclerosis; neoplastic and hyperplastic
 alterations in the central nervous system, also to promote attachment of
 protheses. The present sequence represents a patched protein, which is
 used in the exemplification of the present invention

XX Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 5; Length 1299;

Best Local Similarity 20.4%; Pred. No. 6.7e-33;

Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

QY 321 KGTSLSDKLSFSTHTLLGQFFQGMGTWASWPLTILVLSVIPWALLAAGLVTETLTPDV 380

Db 43 KKGARGRTAIVLRVFSQSHLETSSVQKXGKGVLFVAILVLSFCVGLKSAQHSKVH 102

QY 381 ELWASPNQARSEKAFPHQHFQ-PFFRTNQVILTA---PNRSSYRYSLLGLKPNFSGIL 436

Db 103 QLWIQEGGGLAEALAYTKTKIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153

QY 437 DLDLLELLELLEQLERLHLQVMSPEAQRNLSLODICVAPLNPNTSLY-----DC 486

Db 154 HLEVLVKAIVAKVHLVDYDE-W-----GLRDMCNMPSFPFGIYYIEQILRLHLP 204


```
CC invention.
XX
SQ Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 6; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSKLSFSTHTLLGPOFGQGTWASWPLTILVSLVPVVALAAGLVFTLETDPV 380
DQ 43 KGKARGSRATILRSVFQSHLETLSGSSVQKHAGKLVFVAILVLSTFCVGLSKSAQHSKVH 102
QY 381 ELWASPNQARSKAFHQHFG-PFRTNQVILTA---PNRSSRYDSSLGKPNFSGIL 436
DQ 103 QLMIOEGGGLAEALAYTKTIGEDSATHQLLIQTTHDPNASVLPQALL-----A 153
QY 437 DLDLLELLELOERLRHLOVMSPEAQNRISLODICVAPLNPDTLSY-----DCC 486
DQ 154 HLEVLVATAVKHLVDTE-W-----GLRDMCMNPSTPSEGIYYIEQILRHLPICS 204
QY 487 INSLLOYFQN-----NRTLLILITAN-----QTLMGQTSQVDWKDH 521
DQ 205 IITPLCFWEGSGLLGESAVVTPGLNQLRLTTLNPASVMQYMKQKSEKISPDFETV 264
QY 522 FLYCANAPLTFKDGTTALALSCM-----ADYGAPVPFFLAIG--GYK 560
DQ 265 EQMKRGAIA---SGYMEKPCPLNPLNCPDTPAPNKNSTQPPDVGA---ILSGGCGYA 317
QY 561 GKDYSAEALIMTFSLNNYPAGDPLAQA-----589
DQ 318 AKHMHPBELIVGAKRN-RSGHLRKAQALQSVOQLMTEKMYDQMDNYKVHHLGWTOE 376
QY 590 -----KLWEEAF---LEEMRAFQRRMAGMFQVTFPATERLSLEDEINRTTAEDLIPATSY 640
DQ 377 KAAEVLNAMQNRFSREVEQLLRQSRIATYDIYVFSAAALDDILAKFSPSALSIVIGV 436
QY 641 IVIFLYISALGYSWSRVMDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQ 700
DQ 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGAVGLLMCFSTAGLGLSALLGIVFNAASTQ 492
QY 701 VVPELVLSVADNIFIVLEYQRLPRPGEPREVHIGRALGRVAPSMILCSLSEACFPL 760
DQ 493 VVPELALGVGDVHIFMLTAAYAESNR--EQTKL-----ILKRVGPSILFSACSTAGSFA 546
QY 761 GALTMPAVRTALTSLGLAVLDLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
DQ 547 AAFIPVPAKVFLQAAIVMCSNLAALIVFPAMISLDLRRTAGRADIFCCCFPVWKEQ 606
QY 815 POELPP-----PQGBGL-----LLGPF 832
DQ 607 PKVAPPVLPLNNNGRGARHPKSCNNNRVALPAQNPLLEQRADIPGSSHSLASPSLATPA 666
QY 833 QKAYAPFLHWTTRGVVLLFLAFVLSYSCHISVGLDQELALPKDSYLDYFLFLNR 892
DQ 667 FOHYTFPLMRWVKFLTVMGFLAALISSLYASTRLQDGLDIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAUVVFTVTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSVLAIP--- 949
DQ 727 LP--GFYSYVATQG-----NFEYPTQQQLLDY-HDSFVRVPHVI 764
QY 950 -----ASSWVDDFDIMLTPSSCCRLYISGPNKQKFCPSVNSMLCNLCN----- 993
DQ 765 KNDNGGLPFWLLLFSEWL-----GNLQKIPEEYRDGRLTKECWFPNASSDA 812
QY 994 -----MSITWGSVRFSVEQ-----PHKYL-PWFLNDRNRNICKPKGG 1028
DQ 813 ILAYKLIVGTGHVDNPDVKELVLTNRLVNSDGIINQAFYNILSAW-----ATNASSPTL 868
QY 1029 LAA--YSTSVNLTSQVLASRFMAYHKPKNSQD-----YTEALRAAR 1070
DQ 869 LRANCIRNRANGASQKLYPEPQVHQF--NEYDLKIPKSLPLVVAQMPFYLHGLTDS 926
QY 1071 ELAANITADLRKVPGTDPAFEV--FPYTTITNVFYEQYLITLPEGLFMLSCLVPTFAVNSC 1128
```

```
Db 927 QI-KTLIGHIRDLISVKEGFLPNYPSPGIPFIWEQYMT-LRSSLAMILACVLLAALVLV 984
QY 1129 LLLGLDLRSGLANLLSIVMLIVDTVCGPMALWDISYNVSLINLIVASVAGMSVEFVSHTRS 1188
Db 985 SLLLSVWAALVILSVLASLAQIFGAMTLLGKLSAIPAVILILSVGMMLCF--NVLLIS 1042
QY 1189 FAISTKPTWLERAKETATISMSAVFAGVAMTNLPGLVLGLAKAQIQTQIFPFRINLLITL 1248
Db 1043 LGFMTSVGNRRQVRVOLSMQMSLGPLVHGLMTSGVAVPMLSTSPPEFVIRHFCWLLLVLC 1102
QY 1249 LGLLHGLVFLPVLSVGVDPVNPALAEOKRABEAAVAVMWASCPNHPRSVSTADNIYNV 1308
Db 1103 VGACNSLLVFPILLSVMGPEAE-LVPLE-----HPDRISTPSPLPVR 1143
QY 1309 HS 1310
Db 1144 SS 1145

RESULT 80
ADD25309
ID ADD25309 standard; protein; 1299 AA.
XX AC ADD25309;
XX DT 15-JAN-2004 (first entry)
XX DE Hedgehog associated protein #3.
XX KW hedgehog; patched receptor; spermatogenesis inhibition;
XX KW ovary function inhibition; embryogenesis;
XX KW differential tissue maintenance.
XX OS Unidentified.
XX PN US6576237-B1.
XX PD 10-JUN-2003.
XX PF 16-AUG-2000; 2000US-00639695.
XX PR 30-DEC-1993; 93US-00176427.
XX PR 14-DEC-1994; 94US-00356060.
XX PR 04-MAY-1995; 95US-00435093.
XX PR 05-JUN-1995; 95US-00460900.
XX PA (HARD ) HARVARD COLLEGE.
XX PI (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX PI Ingham PW, McMahon AP, Tabin CJ, Bumcrot DA, Marti-Gorostiza E;
XX WPI; 2003-799823/75.
XX DR N-PSDB; ADD25289.
XX PT Novel isolated antibody which is immunoreactive with a vertebrate
XX PT hedgehog protein sequence that binds with patched receptor, useful for
XX PT blocking action of naturally occurring hedgehog protein, and for
XX PT inhibiting spermatogenesis.
XX PS Disclosure; SEQ ID NO 62; 120pp; English.
XX CC The invention relates to an isolated antibody (I) which is immunoreactive
XX CC with a hedgehog polypeptide (II) that binds to a patched receptor, where
XX CC (II) is encoded by nucleic acid which hybridise to a fully defined
XX CC vertebrate hedgehog (hh) protein. (I) is useful as a hedgehog antagonist
XX CC by blocking action of naturally occurring hedgehog protein, and therefore
XX CC for inhibiting spermatogenesis. (I) is also useful for inhibiting normal
XX CC ovarian function. (I) is useful for blocking the action of one or more
XX CC hedgehog proteins and allows the study of the role of these proteins
XX CC e.g., embryogenesis and/or maintenance of differential tissue. (I) is
XX CC also useful in immunohistochemical staining of tissue samples in order to
XX CC evaluate the abundance and pattern of expression of the hedgehog
```

CC polypeptides. (I) is also useful diagnostically in immunoprecipitation
CC and immunoblotting to detect and evaluate hedgehog protein levels as a
CC part of clinical testing procedure. The present sequence represents the
CC amino acid sequence of a hedgehog associated polypeptide.
XX
SQ Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 7; Length 1299;
Best Local Similarity 20.4%; Pred. NO. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSSTHTLLQFGPGGCTWASNPLTILVSLVPPVVALAAGLVTBLTDPV 380
DB 43 KKGKSGTAIYLRVSFQSHLETGLSSVQKRGKVLFAIIVLSTFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHQHFQ-PPFRTNQVILTA---PNRSSRYDSLLGLPKNFSGIL 436
DB 103 QLMIQEGGLEAEALAYTKTIGEDSATHQLLIQTTHDPAASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLHRHQVSPQARNISLQDICYAPLNPNTSLY-----DCC 486
DB 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCNMPTSPFEGIYYIEQILRHLPSCS 204
QY 487 INSLIQYFQN-----NRTLLLLLTAN-----QTLMGQTSQVDMKDH 521
DB 205 IITPLDCFWEGSQLLGPESAVVPIGLNQRLTWTILNPASVMQMKQKSEKISDFETV 264
QY 522 FLYCANAPLTFKDGTAALSCM-----ADYCAPVFPFLAIG--GYK 560
DB 265 EQMKRAATA---SGYMKPCLNPLNCPDTPAPNKNSTQPPDVG-----ILSGCYGYA 317
QY 561 GKDYSEAEALIMTFSLNYPAGDPLAQA-----589
DB 318 AKHMHWPPELIVGGAKRN-RSGHLKKAQALQSVOQLMTEKEMYDQWQNYKVHILGWTOE 376
QY 590 -----KLWEEAF---LEMBRAFORMAGMFOVTTAERSLDEINRTTAEDLPFATSY 640
DB 377 KAAEVUNAKQRFNSREVEQLLRQSRATNYDIYFSSNAALDDILAKSHPSALSIVIGV 436
QY 641 IVIFLYISALGSSYSWSRVMDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSRLVILQ 700
DB 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNAATQ 492
QY 701 VVPLVLSVGADNIFTVLEYQELPRPCEPREVHIGRALGRVAPSMILCSLSEALCPFL 760
DB 493 VVFFLALGLGVDFHFMLTAAAYAESNR--EQTKL-----ILKKVGSILFSACSTAGSPPA 546
QY 761 GALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKQEAASRLDV-CCCV-----K 814
DB 547 AAFIPVPAKVFCLQAAIVMCSNLAALLVFPAMISLDLRRRTAGRADIFCCCFPWKEQ 606
QY 815 PQBELPP-----PQGEGL-----LLGFF 832
DB 607 PKVAPPVPLNNNGRGARHPKSCNNRVALPAQNPLLEQRADIPGSSHSLSAFLATFA 666
QY 833 QKAYAPELLHWITRGVILLFLALFVGLSYSMCHI SVGLDQELAPKDSVLLDYFLFLNR 892
DB 667 FQHYTFPLMRSWKFLTMVGLAALISSLYASTRLQOGLDIIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVVTTLGYNFSSEAGMNAICSSAGCNFSTQKIQYATEFPPEQSYLAIP--- 949
DB 727 LF--GFYSMAVVTQG-----NFEYPTQQLLRDY-HDSPRVPHVI 764
QY 950 -----ASSWDDFDWLTSPSSCCRLYISGPNKDKFCPSVNSLCKNC-----993
DB 765 KNDNGGLPDPFWLLIFSEWL-----GNLQKIFDEYDRGLTKCEWFPNASSDA 812
QY 994 -----MSITGWSVRPSVEQ-----FKYL-PWFLNDRPNIKCPKGG 1028
DB 813 ILAYKLIVQGHVNDVPDKELVLTNRLVNSDGLINQRAFNYLSAW-----ATWASFTTEL 868
QY 1029 LAA--YSTSVNLTSDDGVQLASRFMAYHKPLKNSQD-----YTEALRAAR 1070

DB 869 LRANCIRNRANGASQGLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQMPPYLHGLTDTTS 926
QY 1071 ELAANITADLRKVPCTDPAFEV--FPVTTNVFVEQVLTILPGLFMLSCLVPTPAVSC 1128
DB 927 QI-KTLIGHIRDLSVKYEGGLPNYPGIPFIWEQYMT-LRSSLAMILACVLLAALVIV 984
QY 1129 LLLGLDLRSGLLMLLSIVMILVDTVGFMALWDISYNVAVSLINLVSAVMSVFEVSHITRS 1188
DB 985 SLLLSWAAVAVILSVLASLAQIFGAMTLLGIKLSAIPAVILILSVGMMLCF--NVLLIS 1042
QY 1189 FAISTKPTWLERAKEATISMSAVFAGVAMTNIPGLIVGLAKAQLIQIFFRNLALLITL 1248
DB 1043 LGFWTSYGNRRVQLSMQSLGLFVHGMLTSGVAVFMTLSTPPEFVIRHFCWLLLVLC 1102
QY 1249 LGLLHGLVFLPVILSVYVGDVNPALALEQKRAEAAVAAVMVASCPNHPSPRVSTADNIYVN 1308
DB 1103 VGACNSLULVPILLUSVMGPEAE-LVPLE-----HPDRISTSPPLVR 1143
QY 1309 HS 1310
DB 1144 SS 1145
RESULT 81
ABW00878
ID ABW00878 standard; protein; 1299 AA.
XX
AC ABW00878;
XX
DT 15-JAN-2004 (first entry)
XX
DE Drosophila patched protein.
XX
KW Cell differentiation; Desert hedgehog; Dhh; Sonic hedgehog; shh;
KW Indian hedgehog; Ihh; skeletogenesis; degenerative disorder; ischaemia;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW Huntington's disease; multiple sclerosis; Pick's disease; aging process;
KW trauma; anoxia; antisense gene therapy; neuroprotective; anticonvulsant;
KW neotropic.
XX
OS Drosophila sp.
XX
FN US2003186357-A1.
XX
PD 02-OCT-2003.
XX
PF 05-JUN-1995; 95US-00462386.
XX
PR 30-DEC-1993; 93US-00176427.
PR 14-DEC-1994; 94US-00356060.
PR 04-MAY-1995; 95US-00435093.
XX
PA (INGH/) INGHAM P W.
PA (MCMA/) MCMAHON A P.
PA (TABI/) TABIN C J.
XX
PI Ingham PW, McMahon AP, Tabin CJ;
XX
DR WPI; 2003-803151/75.
DR N-PSDB; A062116.
XX
PT Modulating cell growth, differentiation or survival, for treating
PT neurodegenerative diseases, such as Alzheimer's or Parkinson's disease,
PT comprises contacting the cell with a hedgehog polypeptide.
XX
PS Disclosure; Page 93-97; Opp; English.
XX
CC The present invention relates to a novel method for modulating growth, the
CC differentiation or survival of a cell. The method involves contacting
CC cell with a hedgehog polypeptide such as Desert hedgehog (Dhh), Sonic
CC hedgehog (shh) and Indian hedgehog (Ihh). The method is used to induce a
CC cell to differentiate to a neuronal cell phenotype. It is used to
CC modulate skeletogenesis. The method is used to treat a degenerative

(HARD) HARVARD COLLEGE.
Ingham PW, McMahon AP, Tabin CJ;
WPI; 2004-794735/78.
N-PSDB; ADU26448.
Modulating growth, differentiation, or survival of a mammalian cell responsive to hedgehog induction, useful for treating e.g. Alzheimer's disease, comprises contacting or treating the cell with a hedgehog polypeptide.
Disclosure; SEQ ID NO 48; 121pp; English.
This invention relates to a novel family of genes referred to as the 'hedgehog' gene family that are involved in pattern formation in embryonic cells. Specifically, it refers to genes that have a broad involvement in the formation and maintenance of ordered spatial arrangements of differentiated tissues in vertebrates. The present invention describes a method for modulating one or more of growth, differentiation or survival of a mammalian cell or an animal responsive to hedgehog induction that comprises contacting or treating the cell with a hedgehog polypeptide that mimics or antagonises the naturally occurring hedgehog protein of the cell. Furthermore, the invention provides methods for inducing a cell to differentiate into a neuronal cell phenotype and also a method of modulating skeletogenesis so as to cause one or both of chondrogenesis and osteogenesis in the target tissue. As such, pharmaceutical compositions derived thereof can be used to treat a degenerative disorder of the nervous system i.e. one that is characterised by neuronal cell death and causes, relative to the absence of hedgehog treatment, prolonged survival of neural cells in the patient. Accordingly, it can be used to treat conditions including neuromuscular disorder, autonomic disorder, central nervous system disorder, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Pick's disease, Huntington's disease and multiple sclerosis. These compositions are shown to exhibit antiparkinsonian, neuroprotective, cerebroprotective, neurotropic, anticonvulsant and vasotropic activities. This polynucleotide sequence is the Drosophila patched protein, a receptor of the hedgehog protein of the invention. NOTE: This sequence is provided in the sequence listing but is not further referred to as specification.
Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 8; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
321 KGTSLSKLSFSTHLLGQFFQGWVWASWPLTILVLSVPPVVALAAGLVFTLTDV 380
43 KKGARSRTAIIYLSVFSQSHLETGSSVQKHAGKVLFAVLVLTSTPCVGLKSAQIHSKVH 102
381 ELWAPNSQARSEKAFHDQHG-PPFRTHQVILTA---PNRSSYRVDLSLLGPKNPSGIL 436
103 QLWIOGGGLEALAYTQIGEDSATHOLLITQTHDPNASVLHPQALL-----A 153
437 DLDLLELELEQLERLHLQVSPQARNISLQDICVAPLNPNTSLY-----DCC 486
154 HLEVLVKAATAVKHLVYDTE-W-----GLRDMCMPTSPFEGIYYIEQILRHLPDS 204
487 INSLILOYFQN-----NRTLILLTAN-----QTLMGQTSQVDWQKH 521
205 IITPLDCFEGSQGLGPESAVVPLGNLQRLTTLNPASVMQYMKQKSEKISPDFETV 264
522 FLYCANAPITFDGTALALSCM-----ADYGAPVPPFLAIG--GYK 560
265 EGYMKRAAIA---SGYMEKPCINPLNPNCPDTPAPKNKSTQPPDVGA---ILSGGCGYGA 317
561 GKDYSBAEALIMTFLSNLYPAGDPLAQA-----589
318 AKHMHWPPELIVGGAQRN-RSHLRKAQALQSVVQLMTEKMYDQWQNDYKVVHHLGWQTE 376
590 -----KLWEEAF---LEEMRAFORRMAGMFQVTFTEARSLEDEINRTTAEDLPFATSY 640

Db 377 KAAEVLNMQRNFSREVEQLLRQSRIATNYDVFSSAALDDILAKFSHPALSIVIGV 436
Qy 641 IVIFLVISIALGSSYSSWSVMVDKATLGLGGVAVVVLGAVMAAGPSPYSIGRSSVLILQ 700
Db 437 ATTVLY---AFTLLRW-RDPVRGSSVGVAGVLLMCFSTAAGLGSALILGIVFNAASTQ 492
Qy 701 VVPFLVLSGADNIFIFVLEYQYRPRRPGEPREVIHGRALGRVAPSMKLSLSEAICFFFL 760
Db 493 VVFLALGLGVDFHIFMLTAAYAESNER--EQTKL-----ILKKVGPISILSACSTAGSFFA 546
Qy 761 GALTMPAPVTRTALTSLGLAVILDFLQMSAFVALLSLSKQBSASLDV-CCCV-----K 814
Db 547 AAAPFVPALKVFCLOAAIYMCNSLAALLVFPFAMISLRLRRRTAGRADIFCCCFPVWKEQ 606
Qy 815 PQELPP-----PGQBSGL-----LLGPP 832
Db 607 PKVAPVPLANNNGEGARHPKSCNNRVLPAPQNPLLEQRADIPGSSHSLSFSLATFA 666
Qy 833 QKAYAPFLHWTIRGVVLLFLALFGVSLYSCHISVGLDQELALPKDSYLDYFLFLNR 892
Db 667 FQHYTPFLMRSMVKPLTVMGFLAALISSLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
Qy 893 YFEVGAPVYVTTLGNFSSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIP--- 949
Db 727 LF--GFYSMYAVTQG-----NFEYPTQQLRDY-HDSFVRVPHVI 764
Qy 950 -----ASSWDDDFIDWLTSPSSCCRLYISGNPKKFCFSTVNSLNLKNC----- 993
Db 765 KNDNGLPDFWLLLFSEWL-----GNLQKIFDEEYRDGRLTKECWFPFNASSDA 812
Qy 994 -----MSITWGSVRPSVEQ-----PHKYL-PWFLNDRPNKCPKGG 1028
Db 813 ILAYKLIQVTHVDNPVDKELVLTNRLVNSDGIINQAFVNYLSAW-----ATWASSPTEL 868
Qy 1029 LAA--YSTSVNLTSDGVLASRPMAYHKPLKNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQOKLYPEPRQVHQD--NEYDLKIPKSLPLVYAQMPFYHLGLTDT 926
Qy 1071 ELAANITADLRKVPCTDPAFEV--PFTYITNVFYEQYLTILPEGLPMLSLCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDLRSVKYEGFGLPNYPSPGIPPIFWEQYMT-LRSSLAMILACVLLAALVLV 984
Qy 1129 LLLGLDLRSGLNLLSIVMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHTRS 1188
Db 985 SLLLSVAAVILVILSVLASLAQIFGAWTLGLKLSAIPAVILILSVGMULCF--NVLS 1042
Qy 1189 FAISTKPTWLERAKATISMSGSAVPAVAMTNLPGILVLGLAKAQIIOIFFRLNLLITL 1248
Db 1043 LGFMTSVGNRRQRVQLSMQMSLGPLVHGMLTSGVAVFMLSTSPFVIRHFCWLLILVLC 1102
Qy 1249 LGLLHGLVPLVILSVGPDVNPALALEQKRAEBAVAAVMVASCNHPSEVSTADNIYN 1308
Db 1103 VGACNSLIVFPILLSMVGGEAB-LVPLE-----HPDRISTPSPLPVR 1143
Qy 1309 HS 1310
Db 1144 SS 1145
RESULT 86
ADW25903
ID ADW25903 standard; protein; 1299 AA.
XX
AC ADW25903;
XX
DT 24-MAR-2005 (first entry)
XX
DE Drosophila sp. patched protein.
XX
KW Differentiation; cell growth; pharmaceutical; patched protein.
XX
OS Drosophila sp.

XX Key Location/Qualifiers
PH 93..426
FT Domain /note = Extracellular domain
FT Domain 700..966
FT /note = Extracellular domain
XX US2005004028-A1.
XX PN
XX PD
XX PF
XX PR 28-APR-2004; 2004US-00835517.
XX PR 30-DEC-1993; 93US-00176427.
XX PR 14-DEC-1994; 94US-00356060.
XX PR 04-MAY-1995; 95US-00435093.
XX PR 05-JUN-1995; 95US-00462386.
XX PR 20-OCT-1997; 97US-00954771.
XX (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX PA
XX PI Ingham PW, McMahon AP, Tabin CJ;
XX WPI; 2005-065255/07.
DR N-PSDB; ADW25897.
XX
XX Use of hedgehog polypeptide for modulating growth, differentiation, or
PT survival of a cell; inducing a cell to differentiate to a neuronal cell
PT phenotype; or treating a degenerative disorder, e.g. Alzheimer's disease.
XX
XX Disclosure; SEQ ID NO 48; 122pp; English.
XX PS
XX The present invention relates to the hedgehog proteins and their
CC polynucleotides. The invention is useful for modulating growth,
CC differentiation and survival of a cell, inducing a cell to differentiate
CC to a neuronal cell phenotype and modulating skeletogenesis. The invention
CC is also useful for treating degenerative disorders of the nervous system
CC like Alzheimer's disease, Parkinson's disease, amyotrophic lateral
CC sclerosis, Pick's disease, Huntington's disease and multiple sclerosis,
CC neuronal damage resulting from anoxia-ischemia and trauma, neuronal
CC degeneration associated with a natural aging process and neuronal cell
CC death. The present sequence is the Drosophila sp. patched protein. This
CC sequence is used to illustrate the activity of vertebrate embryonic
CC pattern-inducing proteins.
XX
SQ Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 9; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSFSTHTLLGQFFQGWGWASWPLTILVSVIPVVALAAGLVFTELTPDV 380
DB 43 KGARGSRTHYLRVFSQSHLETIGSSVQXGACKVLFVAILVLTSTFCVGLKSAQHSKH 102
QY 381 ELMSAPNSQARSEKAFDHFQGFPPFRNTQVILTA---PWRSSRYVDSLLGKPNFSGIL 436
DB 103 QLMTQEGGGLAEALAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLHLQWSPAQRNISLQDICYAPLNPNTSLY-----DCC 486
DB 154 HLEVLKATAVKVLHYDTE-W-----GLRDMCNMPTSPSEGIYIEQILRHLPIS 204
QY 487 INSLQLQVFQ-----NRTLLLTAN-----QTLMGQTSQVDWKDH 521
DB 205 IITPLDCFWESQLLGPESAVVIFGLNQLRLTTLNPAVQYMKQWSEKISFDPETV 264
QY 522 FLYCANAPLTKDGTALALSCM-----ADYGAPVFFFLAIG--GYK 560
DB 265 EQYMKRAAIA---SGYMEKPCLNPLNCPDPAFNKNSQPPDVGA-----ILSGGCGYGA 317
QY 561 GKDYSEAEALIMTSLNNYPAGDPLRAQA----- 589

DB 318 AKTHWPEELIVCGAKRN-RSGHLRKAQALQSVVQLMTEKEMYDQWQDNKYKVHHLGWTOE 376
QY 590 -----KLWEEAF---LEEMRAFQRMAGMFOVFTFAERSLEDEINRTTAEDLPIPATSY 640
DB 377 KAAEVLNMQNFRSREVEQLLRQSRITATYDIYVFFSSAALDDILAKFSPHSALSIVIGV 436
QY 641 IVIFLYISLALGSYSSWSRUMVDSKATLGLGGVAVVLGAVMAAMGPFYSYIGIRSSVLILQ 700
DB 437 AVTVLY---AFCTLLRW-RDPVVGQSQSVGVVLLMCFSSTAAGLSALLGIVFNAASTQ 492
QY 701 VVPFLVLSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMCLLSLSAICFFL 760
DB 493 VVPFLALGLGVDFHIFMLTAAYAESNR--EQTKL-----ILKVGSPSLFSACSTAGSFFA 546
QY 761 GALTMPAVRTFALTSGLAVIDFLLOMSAFVALLSLDSKQEQASRLDV-CCCV-----K 814
DB 547 AAFITVPALKVFCLOAAIVMCSNLAALVFPAMISLDLRRRTAGRADIPCCCFPVWKEQ 606
QY 815 PQELPP-----PGQGEGL-----LIGFF 832
DB 607 PKVAPPVPLNNNGRGARHPKSCNNRVALPAQNPLEQRADIPGSSHSLSASFATFA 666
QY 833 QKAYAPFLHWTIRGVVLLFLALFGVSLYSMCHISVGLQDELALPKDSYLLDFLFLNR 892
DB 667 FQHYTPFLMRSWKFLTMGFLAALISSLYASTRLQDGLDIDLVPKDSNEHFKFLDAQTR 726
QY 893 YFEVGAPVYFTTLGYNFSSEAGNNAICSSAGCANNFTQKIOYATPEPESYLAIP--- 949
DB 727 LF--GFYSMAVATQ-----NFETPTQQQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWVDDFIDLWTPSSCCRLYISGPNKDKFCPTVNSLNLKNC----- 993
DB 765 KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYVRDGLTKECWFPNASSDA 812
QY 994 -----MSTMGSVRPSVEQ-----FKYL-PWFLNDRPNKCPKGG 1028
DB 813 ILAYKLI VQTGHVNDPVDKELVLTNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTL 868
QY 1029 LAA--YSTSVMLTSDGOVLASRFMAYHKLKNSQ-----YTEALRAAR 1070
DB 869 LRANCIRNRANGASQGLYPBPQYFHPQ---NEYDLKIPKSLPLVYAQMPPFLHGLTDS 926
QY 1071 ELAAMITADLRKVPCTDPAFEV--PPYTTITNVFVEQYLTILPEGLFMSLCIVPTFAVSC 1128
DB 927 QI-KTLGHIRDLSVKEGFGLPNYPSPGIPPIFWEQYMT-LRSSLAMILACVLLAALVLV 984
QY 1129 LLLGLDLRSLNLLSIVMILVDTVGFMAALWDISYNAVSLINLVSAGVMSVEFVSHITS 1188
DB 985 SLLLSVMAAVILSVLASLAQIFGAMTLLGILKLSAIPAVILISVGMMLCF--NVLLIS 1042
QY 1189 PAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQILQIFPFLMLLITL 1248
DB 1043 LGFMTSVGNRQRRVQLSMQSLGPLVHGMLTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102
QY 1249 LGLHGLVPLPVLISVGVDPVNPALAEQKRAEEAAVAVMVASCNHPERSVSTADNIYN 1308
DB 1103 VGACNSLLVFPILLSWGEAB-LVPLE-----HPDRISTPSPLPVR 1143
QY 1309 HS 1310
DB 1144 SS 1145
RESULT 87
AEB22486
ID AEB22486 standard; protein; 1299 AA.
XX
AC AEB22486;
XX
DT 08-SEP-2005 (first entry)
XX
DE Fruit fly patched protein.

XX Patched protein; Antidepressant; Tranquillizer; Nootropic;
 KW Neuroprotective; neuroleptic; therapy; major depressive disorder;
 KW cognitive disorder; learning disorder; dementia;
 KW attention deficit hyperactivity disorder; psychiatric disorder;
 KW anxiety disorder; panic disorder; obsessive-compulsive disorder; phobia;
 KW post traumatic stress disorder.
 XX
 OS Drosophila melanogaster.
 XX
 XX WO2005061002-A2.
 XX
 XX 07-JUL-2005.
 XX
 XX 15-DEC-2004; 2004WO-US042271.
 XX
 XX 19-DEC-2003; 2003US-0531201P.
 XX
 XX (CUR1-) CURIS INC.
 XX
 XX Rubin IL;
 XX
 XX WPI; 2005-522111/53.
 XX N-PSDB; AEB22485.
 XX
 XX Treating depression, involves administering to patient exhibiting
 PT symptoms of depression agonist of hedgehog signaling to ameliorating some
 PT or all of symptoms of depression.
 XX
 XX Disclosure; SEQ ID NO 23; 241pp; English.
 XX
 CC The invention relates to treating (M1) depression, prophylactic treatment
 CC (M2) of a patient at risk of developing depression, enhancing (M3)
 CC cognitive function in a mammal, prophylactic treatment (M4) of a patient
 CC at risk of developing a deficiency in cognition, enhancing (M5) memory in
 CC a mammal, prophylactic treatment (M6) of a patient at risk of developing
 CC a deficiency in memory, treating (M7) an emotional disorder characterized
 CC by abnormal activity of a central nervous system (CNS) of a mammal,
 CC prophylactic treatment (M8) of an emotional disorder characterized by
 CC abnormal activity of a CNS of a mammal, treating (M9) non-Alzheimer's
 CC dementia or attention deficit hyperactivity disorder (ADHD) or
 CC prophylactic treatment (M10) of non-Alzheimer's disease or ADHD, involves
 CC administering an agonist of hedgehog (Hh) signaling to the patient or
 CC mammal, thus ameliorating some or all of the symptoms of depression,
 CC preventing deficiency in cognitive function, enhancing memory in the
 CC mammal, preventing memory deficiency, treating the emotional disorder,
 CC preventing the emotional disorder, treating the behavioral disorder or
 CC preventing the behavioral disorder. (M1) is useful for treating
 CC depression. (M2) is useful for prophylactic treatment of a patient at
 CC risk of developing depression. (M3) is useful for enhancing cognitive
 CC function in a mammal. (M4) is useful for prophylactic treatment of a
 CC patient at risk of developing a deficiency in cognition. (M5) is useful
 CC for enhancing memory in a mammal. (M6) is useful for prophylactic
 CC treatment of a patient at risk of developing a deficiency in memory. (M7)
 CC is useful for treating an emotional disorder characterized by abnormal
 CC activity of a CNS of a mammal. (M8) is useful for prophylactic treatment
 CC of an emotional disorder characterized by abnormal activity of a CNS of a
 CC mammal. (M9) is useful for treating non-Alzheimer's dementia or ADHD.
 CC (M10) is useful for prophylactic treatment of non-Alzheimer's disease or
 CC ADHD. The emotional disorder is depression, anxiety disorder, panic
 CC disorder, obsessive compulsive disorders, social anxiety/phobic disorder,
 CC or posttraumatic stress syndrome. The present sequence represents a fruit
 CC fly patched protein, a protein known to interact with hedgehog proteins.
 CC NOTE: AEB22485 is shown as encoding the protein appearing as AEB22486 in
 CC the sequence listing, however AEB22486 has not been given a separate SEQ
 CC ID NO and therefore both have termed SEQ ID NO 23.
 XX
 XX Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 9; Length 1299;

Best Local Similarity 20.4%; Pred. No. 6.7e-33;

Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

QY 321 KGTSLSDKLSFSTHTLLGQFFQGMGTWASWPLTILVLSVIPVVALAAGLVFTELTDDV 380
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 43 KKGARGRTAIYLRVSFQSHLETLGSSVQKHAGKVLFAVLVSTFCVGLKQAHSKVH 102
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 381 ELWASNPQARSEKAPHQHQFQ-PPFRTNOVILTA---PNRSSYRVDLSLLGPKNPSGIL 436
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 103 QWIOEGGGLEALAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 437 DIDLLELLELQERLRLHLQVMSPEAQRNISILODICIYAPLNPNTSLY-----DCC 486
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 154 HLEVLVATAVKAVHLYDTE-W-----GLRDMCNMPESTPSPGEGIYYIEQILRHLP 204
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 487 INSLLOYFQN-----NRTLLLLLTAN-----QTLMGQTSQVDMKDH 521
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 205 IITPLDCFMEGSQLLGPSAVVIPGLNORLLTTLNPASVMQYMKMKSEKISIPDET 264
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 522 FLYCANAPLTPKDGTAALASCM-----ADYGAPVPPPLAIG--GYK 560
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 265 EYMKRAATA---SGYMEKPCNLNPNPCPDTPAPNKNSTQPPDVGA-----ILSGG 317
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 561 GRDYSEAEALIMTFSLNYPAGDPRLAQA-----589
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 318 AKCHMWPEELIVGGAKRN-RSCHLRKAQALQSVVQLMTEKEMYDQMDNYKVHHLGWTQE 376
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 590 -----KLWEAP-----LEEMAPQRMAGMFOVTTAERSLSEINRTTAEDLPATSY 640
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 377 KAAEVLNAWQRNFSREVQLLRKQSRITATNYIYVFSAAALDDILAKFSPSALSIVIG 436
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 641 IVPELYISALGSSYMSRVMYDSKATGLGQVAVVLGAVMAAMGFFSVGLSRSSLVILQ 700
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGAVGLMCFSTAAGLGLSALLGVFNAASTQ 492
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 701 VVPFVLVSGADNIFVLEYQRLPRRPGEPREVIHGRVAPSMILCSLSEATCFPL 760
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 493 VVPFALGLGVDDHIFMLTAAYAESNR--EQTKL---ILKKVGSILFSACSTAGSFFA 546
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 761 GALTMPAVRTFALTSGILAVLDLFLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 547 AAFIVPALKVFCLOAAIVMCSNLAALLVFPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 815 PQELPP-----PQCGSL-----LLGFP 832
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 607 PKVAPFVLPLNNGRGARHPKSCNNRVVALPAQNPLLEQRADIPGSSHSLSFSLATFA 666
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 833 OKAYAPFLHMTIRGVLWLLFLALPGVSLYSMCHISVGLDQBLAPKQSYLLDYFLFLNR 892
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 667 FQHYTFPLMRSVKELTWGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 893 YPEVGAPYVFTTLGYNFSSEAGMNAICSSAGCNPSPFTQIQTAYATEPPEQSYLAIP--- 949
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 727 LF--GFYSMYAVTQG-----NFEVPTQQLRDY-HDSFVRVPHVI 764
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 950 -----ASSWDDDFIDWLTSSCCRLYISGNPKDKFCBSTVNSLNLKNC----- 993
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 765 KNDNGGLPDPFLLLPSEWL-----GNLQKIFDEYDRDGLRTKECFWPNASSDA 812
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 994 -----MSITMGSVRSVBO-----FHKYL-PWFLNDRDNKCPKGG 1028
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 813 ILAYKLIYQTEGVDPNVDKELVLTNRLVNSDGIINQRAFYNYLSAW-----ATNASPT 868
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 1029 LAA--YSTSVNLTSQGVLAGRFMAYHKPLKNSQD-----YTEALRAAR 1070
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 869 LRANCIRNRANGASQGLVPBPQYFHPQ--NEYDLKIPKSLPLVYAQMPFYLHGLTOTS 926
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 1071 ELAANITADLRKVPQTDPAFEV--PPYITVVFYEQLTILPEGLFMLSCLVPTFAVNSC 1128
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 927 QI-KTLIGHIRDLVSVKYEGFLPNYPGSGIPFIFWEQYMT-LRSSLAMTLCVLLAALVL 984
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 1129 LLLGLDLRLSGLLNLISIVMLVDTVGFMAWDISVNAVSLINLVSAGVSVEFVSHITRS 1188
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 985 SLLLSVWAVALVILSVLASLAQIFGAMTILGILKLSAPVAVLILSVGMKLCF--NLVIS 1042
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 1189 FAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLITL 1248
 DB ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :


```

Db 960 ACVLLAALVLSVLLLSVAAVLLVLSVLAASQAIFGAMTLTGKLSAIPAVILSVGM 1019
Qy 1178* SVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQI 1237
Db 1020 MLCF--NVLISLGFMTSVGNRRVQLSMQMSIGPLVHGLTSGVAVFMLSPPFEFVIR 1077
Qy 1238 PFRLNLLTLGLLGLVFLPVLISYVGPDPVNPALLAKOKRAEAAVAVMVASCPNHP 1297
Db 1078 HFCWLLLVLCVGCACNSLLVFPILLSVMGPEAE-LVPLE-----HPD 1118
Qy 1298 RVSTADNIYVNH 1310
Db 1119 RISTPSLPVRSS 1131

RESULT 89
AAB67157
ID AAB67157 standard; protein; 1285 AA.
XX
AC AAB67157;
XX
DT 12-APR-2001 (first entry)
XX
DE Fruit fly patched protein.
XX
KW Fruit fly; patched; PTC; segment polarity; limb patterning; development;
KW hedgehog; antibody; human; mouse.
XX
OS Drosophila melanogaster.
XX
PN US6172200-B1.
XX
PD 09-JAN-2001.
XX
PF 20-OCT-1997; 97US-00954668.
XX
PR 07-OCT-1994; 94US-00319745.
XX
PR 06-OCT-1995; 95US-00540405.
XX
PA (STRD ) UNIV LELAND S STANFORD.
XX
PI Scott MP, Goodrich LV, Johnson RL;
XX
DR WPI; 2001-136884/14.
XX
DR N-PSDB; AAF32178.
XX
PT Novel monoclonal antibody useful in diagnostic assays for detection of
PT presence of protein on surface of cells specifically binds to naturally
PT occurring patched protein, other than Drosophila patched protein.
XX
PS Disclosure; Col 35-42; 39pp; English.
XX
CC The present invention provides a monoclonal antibody which specifically
CC binds to a patched protein (PTC) other than that from Drosophila. Also
CC given are the protein and coding sequences of patched from the beetle,
CC mouse, butterfly, fruit fly, mosquito and human. Patched is a segment
CC polarity gene involved in limb patterning. The sequences can be used to
CC study development and to isolate the patched ligand, hedgehog. In
CC addition, antibodies can be used to detect the PTC protein on cell
CC surfaces or to inhibit the transduction of signal by the PTC ligand by
CC competing for its binding site
XX
SQ Sequence 1285 AA;

Query Match 6.3%; Score 438.5; DB 4; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.8e-30;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

Qy 321 KGTSLSDKLSFSTHTLLGCGFWGCTWVASWPLTILVLSVIPVVALAAGLVPTLTDVPV 380
Db 43 KKGKAGSRITVLRVSFQSHLETGSSVQKAGKRVLFVAILVLSTFCVGLKSAQIHSKVH 102

```

```

Qy 381 ELWAPNSQARSEKAPHDQHG-PPERTNQVILTA---PNRSSRYRYSLLGPKNFSGIL 436
Db 103 QLWIOGEGSLEAELAVTKTIGEDSATHQLLIQTTHDFNASVLHPQALL-----A 153
Qy 437 DLDLLELLELQERLHLQVMSPEAQRNLSIODICVAPLNPNTSYL-----DCC 486
Db 154 HUEVLVKATAVKVHLYDTE-W-----GLRDMCNMPTSPFEGIYYIEQILRHLIPCS 204
Qy 487 INSLIQYFON-----NRTLILLLTAN-----QTLMGQTSQVDPKDH 521
Db 205 IITPLDCFWEGSQLGPESAVVIPGLNQRLMTTLNPASVMQMKMKSEKISFPFETV 264
Qy 522 FLY-----C-----ANAPLTFKD-----GTALASCADYAGAPVPPP 553
Db 265 EOYMKRAAIGSGYMEKPCINPLNPNCPDTAPNKNSTQPPDVGAIISSGCGVG-YAAKGMHW 323
Qy 554 ---LAIGYK---GKDYSEABALIMTFSL-----NNYPAGDPRLAQAK-----590
Db 324 PEELIVGGRKRNRSGHLRKAQALQSVVQLMTSEKEMVQMDNYKVHHLGWTQSKAAEVLN 383
Qy 591 LWEEAF---LEEMRAFORRMAGMFOVTFPAERSLEDEINRTTAEIDLPIFATSVIVFLYI 647
Db 384 AWQNPFSREVEQLLRKQSRATNYDIYVFPSSAALDDILAKFSPSALSIVIGVAVTVLY- 442
Qy 648 SLALGSYSWSRMVDSKATLGLGGVAVVLGAVMAAMGFFSYLGI-----692
Db 443 ---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
Qy 693 -RSSLVI---LQVVPFLVSVGADNIFIFVLEYORLPRRPGEPREVIHGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVVPLALGLGVDFHFI-----VGPSTI 533
Qy 748 LLCSLSEACPFELGALTMPAVRTFALTSGLAVIDELLQMSAFVALLSLDSKROASRL 807
Db 534 LFSACSTAGSFFAAAFIPVPALKVFCLOAIVMCSNLAAALLVFPAMISLDRRTAGRA 593
Qy 808 DV-CCCV-----KQOELPP-----PGQ 823
Db 594 DIFCCCFPMWKEQPKVAPPVPLPLNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGS 653
Qy 824 GEGL-----LLGFFQKAYAPFLHWTIRGVVLLLPALFGVLSYSMCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFQHYTPFLMRSWVKFLTWGFLAALISLSYASTRLQDGLDII DLVPK 713
Qy 880 DSYLLDYFLFLNRYFEVGAPVYVTTGLVNPFSSEAGMNAICSSAGCNCNPFQKIQAYTE 939
Db 714 DSNKHFIDAQTRLP--GFYSWIAVTQG-----NFEYPTQOQLLRD 752
Qy 940 FPEQSYLAIP-----ASSWYDDFIDMLTPSSCCRLYISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNGGLPDPFWLLLFSEWL-----GNLQKIFDEYRDGRL 798
Qy 990 LKNC-----MSITMGSVRPSVBO-----PHKYL-PWF 1015
Db 799 TKECWFFPNASSDAIILAYKLIIVQTGHVDNPNVDKELVLTNRLVNSDGIINQRAFYNLSAWA 858
Qy 1016 LNDRENIKCPKGLAAYSTSVNLTSDGVLASRFMAVHKPLKNSOD-----1061
Db 859 TND-----VPAIG-----ASQKGLYPEPROYFHP--NEYDLKIPKSLPLVYAQM 901
Qy 1062 --YTEALRAARELANITADLRKVPGTDPAPEV--FPYITNVFYEQVLTILPEGLFMLS 1117
Db 902 PPLYHGLTDTSQI-KTLIGHIRDLRSVKYEGFGLPNYPGSIPIPFWEQYWT-LRSSLAMTL 959
Qy 1118 LCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAWDISYNAVSLINLSAVGM 1177
Db 960 ACVLIAALVLSLILLSVAAVLSVLAQAIFGAMTLTGKLSAIPAVILSVGM 1019
Qy 1178 SVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQI 1237
Db 1020 MLCF--NVLISLGFMTSVGNRRVQLSMQMSIGPLVHGLTSGVAVFMLSPPFEFVIR 1077
Qy 1238 PFRLNLLTLGLLGLVFLPVLISYVGPDPVNPALLAKOKRAEAAVAVMVASCPNHP 1297
Db 1078 HFCWLLLVLCVGCACNSLLVFPILLSVMGPEAE-LVPLE-----HPD 1118
Qy 1298 RVSTADNIYVNH 1310
Db 1119 RISTPSLPVRSS 1131

```

Db	1078	HFCMLLVLCVGCAGNSILVFPILLSMVGPEAE-LVPLE-----	HPD 1118
Qy	1298	RVSTADNIYNVNH5 1310	
Db	1119	RISTPSPLPVRSS 1131	
RESULT 90			
AAG79574			
ID	AAG79574	standard; protein; 1285 AA.	
XX	AC		
XX	AAG79574;		
XX	23-DEC-2002	(first entry)	
XX	DE	Drosophila patched protein.	
XX	KW	Patched; ptc; basal cell nevus syndrome; BCNS; wound healing;	
XX	KW	developmental disorder; spina bifida; craniofacial abnormality;	
XX	KW	basal cell carcinoma; skin; brain tumour; transgenic; mouse;	
XX	KW	proliferation; oncogenesis; signal transduction; cancer; aging.	
XX	OS	Drosophila melanogaster.	
XX	XX	Location/Qualifiers	
XX	FT	Misc-difference 758	
XX	FT	/note= "Encoded by GTGCGG"	
XX	FT	Misc-difference 812	
XX	FT	/note= "Encoded by ATG"	
XX	US64293354-B1.		
XX	PN		
XX	XX		
XX	PD		
XX	XX		
XX	PF	22-AUG-1997; 97US-00918658.	
XX	XX		
XX	PR	07-OCT-1994; 94US-00319745.	
XX	PR	06-OCT-1995; 95US-00540406.	
XX	PR	31-MAY-1996; 96US-00656055.	
XX	XX	(STRD) UNIV LELAND STANFORD JUNIOR.	
XX	PA	(REGC) UNIV CALIFORNIA.	
XX	XX		
XX	PI	Scott MP, Goodrich LV, Johnson RL, Epstein E;	
XX	XX		
XX	DR	WPI; 2002-722086/78.	
XX	DR	N-PSDB; ABA00377.	
XX	XX		
XX	PT	Transgenic mouse, useful for cancer drug screening, comprises disruption	
XX	PT	of allele of patched locus which predisposes the mouse to develop a	
XX	PT	proliferative disorder.	
XX	XX		
XX	PS	Example; Col 57-64; 58pp; English.	
XX	XX		
XX	CC	This sequence shows drosophila patched (ptc) protein. Mutations in the	
XX	CC	patched gene are characteristic in basal cell nevus syndrome (BCNS), an	
XX	CC	inherited disorder with an increased risk of developmental disorders such	
XX	CC	as spina bifida and craniofacial abnormalities, basal cell carcinoma of	
XX	CC	the skin and brain tumours. The human and mouse ptc sequences show 86%	
XX	CC	homology. The transgenic mouse of the invention has a genome comprising a	
XX	CC	disruption of at least one allele of a patched locus, where the	
XX	CC	disruption predisposes the mouse to develop a proliferative disorder. The	
XX	CC	transgenic mouse is useful for studying patched function and regulation,	
XX	CC	for e.g. a series of small deletions and/or substitutions made in the	
XX	CC	patched gene, to determine the role of different exons in oncogenesis,	
XX	CC	signal transduction, etc; and for drug screening which is useful for	
XX	CC	treating cancer or developmental abnormalities attributed to a defect in	
XX	CC	patched function such as wound healing and aging	
XX	SQ	Sequence 1285 AA;	
Query Match 6.3%; Score 438.5; DB 5; Length 1285;			
Best Local Similarity 20.0%; Pred. No. 1.8e-30;			
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;			
Qy	321	KGTSLSDKLSFSTHTLLGQFFGQGWTVASWPLTILVSVIPVVALAAGVFTTELTDVPV 380	
Db	43	KGKARGRTAIYLRVSFQSHLETGLSSVQKHAGKVLFAVAILVSTFCVGLKSAQIHSKVH 102	
Qy	381	ELWSAPNSQARSEKAFHDQHFQ-PPFRTNQVILTA---PNRSSYRYDSLLLGPKNFSGIL 436	
Db	103	QLMIQGGRLAEALAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153	
Qy	437	DLDLLLELLEQLERLRLHQWSPEAQRNISLQDICYAPLNPDNTSLY-----DCC 486	
Db	154	HLEVVKATAVKVHLYDTB-W-----GLRDMCMNPSTPSPEGIYIIEQILRHLPSCS 204	
Qy	487	INSLLYQFQN-----NRTLILLTAN-----OTLMQSTQSDVQDKH 521	
Db	205	IITPLDCFWEGSQLGPESAVIPGLNQLLLTLLNPASVMQMKQMSSEKISFDETV 264	
Qy	522	FLY-----C-----ANAPLTPKD-----GTALALSCMADYGAIVPFF 553	
Db	265	EQMKRAAIGSGYMEKPCPLNPNPCPDTPAPKNKSTOPDPVGAILSGCYG-YAAKMHW 323	
Qy	554	---LAIGGYK---GKDYSEAEALIMTSL-----NNYPAGDPRLAQA-K----- 590	
Db	324	PEELIVGGRKRNRSGLRKAQALQSVVQLMTEKEMYDQMDNYKVHHLGWTQEKAAEVLN 383	
Qy	591	LWEEAF---LEEMRAFORRMAGMFQVTTFAERSIESTEINETTAEEDLPFATSYIVFLYI 647	
Db	384	AWQRNFSREVEQLLRKQSRATNYDIYVFSAAUDDILAKFSPSALSIVIGNAVTVLY- 442	
Qy	648	SLALGSYSSWSRVNVDKATLGLGGVAVVLGAVMAAMGFFSYLGI----- 692	
Db	443	--AFTLLRW-RDPVRCGSSVGAGVLLMCFSTAAGLGSALLGIVFNALTAAYAESNRR 499	
Qy	693	-RSSLVI-----LQVPFLVLSVGADNIFIFVLEYQRLPRPGEPREVHIGALGRVAPSM 747	
Db	500	EQTKLILKNASTQVVPFLALGLGVDHIF1-----VGPSI 533	
Qy	748	LLCSLSEAI CFFLGALTPMPAVRTFALTGLAVILDFLLOMSAFVALLSLDSKRQEASRL 807	
Db	534	LFSACSTAGSPFAAATPVPALKVFCLOAAIWCNSLAAALLVFPAMISLDRRTAGRA 593	
Qy	808	DV-CCCV-----KPQELPP-----PGQ 823	
Db	594	DIFCCCFPVWKEQPKVAPPVPLNANNNGRGARHPKSCNNNRVPLPAQNPLEORADIPGS 653	
Qy	824	GEGL-----LLGFFOKAYAPFLHLHWITRGVLLLLFLALFGVLSYNSCHISVGLDQELAPK 879	
Db	654	SHSLASFSLATFAFQHYTFFLMRSWVKFLTYMGFLAALISSLYASTRLQDGLDIIDLVPK 713	
Qy	880	DSYLLDYFLNRYFEVGPVVFVTTLGYNFSSEAGMNAICSSAGCANNFSFTQKIQYATE 939	
Db	714	DSNEHKFLDAQTRLF--GFYSMYATQG-----NFEYPTQQLLRD 752	
Qy	940	FQESYLAIIP-----ASSWDDFDIMLTPSSCCRLYISGKNKDKFCBSTVNSLNC 989	
Db	753	YHDS--FRVPHVYKNDNGGLPDFWLLLFSEWL-----GNLQKIFDEEYRDL 798	
Qy	990	LKNC-----MSITMGSVRPSVEQ-----PHKYL-PWF 1015	
Db	799	TKECWFPNASSDAILAYKLI VQTGHVDPVDKELVLTNRLVNSDGIINQRAFNYLSAWA 858	
Qy	1016	LNDNRNIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQD----- 1061	
Db	859	TND-----VFAYG-----ASQGLKYPPEQYFHPQ--NEYDLKIPKSLPLVYAQM 901	
Qy	1062	--YTEALRAARELAANITADLRKVPCTDPAFEV--FPYITNVFQYIETILPEGLFMLS 1117	
Db	902	PFLYHGLTDTSQI-KTLIGHIRDLSVKYEGFLPNYPSGIPFIFWEQYMT-LRSSLAMIL 959	
Qy	1118	LCLVPTFAVSCLLLGDLRSGLNLLSIYMIIVDVTGVFMALWDISYNAVSLNLSVAVGM 1177	

Db 960 ACVLLAALVLSLLLSVAAVILSVLASLAQIFGAMTLLGIKLSAIPAVILLSVGM 1019
 QY 1178 SVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQI 1237
 Db 1020 MLCP--NVLLISLGFMTSVGNRRQVQLSMQMSLGPLVHGLTSGVAVFMLSPPPEFVIR 1077
 QY 1238 FFPRNLNLLITLGLLHGLVFLPVILSYVCPDYNPALALEQKRAEAAVAVMVASCPNHP 1297
 Db 1078 HFCWLLLVLCVGCNLSLFFILLSMVGPEAE-LVPLE-----HPD 1118
 QY 1298 RVSTADNIYVNH 1310
 Db 1119 RISTSPPLVRSS 1131

RESULT 91

ABU62150
 ID ABU62150 standard; protein; 1285 AA.

AC ABU62150;

XX 25-AUG-2003. (first entry)

XX Fruit fly patched gene PTC product.

XX Fruit fly; patched gene; PTC; hedgehog protein; gene therapy.

XX Drosophila melanogaster.

XX US2003032085-A1.

XX 13-FEB-2003.

XX 20-OCT-1997; 97US-00954701.

XX 07-OCT-1994; 94US-00319745.

XX 06-OCT-1995; 95US-00540406.

XX (SCOT/) SCOTT M P.

XX (GOOD/) GOODRICH L V.

XX (JOHN/) JOHNSON R L.

XX Scott MP, Goodrich LV, Johnson RL;

XX WPI; 2003-492065/46.

XX New DNA sequence other than present in a chromosome encoding patched gene
 PT other than Drosophila patched gene, useful for preparing transgenic
 PT laboratory animals and to knock out patched protein in embryonic stem
 PT cells.

XX Disclosure; Page 8-10; 40pp; English.

PS The invention relates to a DNA sequence other than present in chromosome
 CC encoding a patched (ptc) gene other than Drosophila patched gene. A cell
 CC expressing the DNA sequence is useful for producing patched protein. A cell
 CC growing the cell expressing the DNA sequence, where the patched protein
 CC is expressed and isolating the patched protein free of other proteins.
 CC The cell expressing the DNA sequence is also useful for screening
 CC candidate compounds for binding affinity to the patched protein, by
 CC combining the candidate protein with the cell expressing the DNA sequence
 CC where the DNA sequence comprises the entire coding sequence under the
 CC transcriptional regulation of the transcriptional initiation region and a
 CC transcriptional termination region functional in the cell, expressing the
 CC patched protein in the cell and assaying for the binding of the candidate
 CC compound to the patched protein. The above cell is useful for screening
 CC candidate compounds for agonist activity with the patched protein. The
 CC DNA sequence is useful for producing all or portions of the patched
 CC protein, as probes for research, diagnosis, binding of hedgehog protein
 CC for its isolation and purification and in gene therapy. The DNA sequence
 CC is also useful as primers for investigating other species and for
 CC isolating genes from various mammalian sources of interest, particularly
 CC from humans or from domestic animals. The DNA sequence is further useful

CC for preparing transgenic laboratory animals and to knock out the PTC
 CC protein in the embryonic stem cells, so as to produce hosts with single
 CC functional patched gene. The present sequence represents the amino acid
 CC sequence of the fruit fly patched gene PTC product
 XX
 SQ Sequence 1285 AA;

Query Match 6.3%; Score 438.5; DB 7; Length 1285;
 Best Local Similarity 20.0%; Pred. No. 1.8e-30;
 Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KGTSLDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLVSVIPVVALAAGLVFTELTTDPV 380
 Db 43 KKGARSGRTAIYLRVSFQSHLETGSSVQKHAGKLVFVAILVLSTFCVLGSAQIHSKVH 102
 QY 381 ELWASPNQARSEKAFHQHFG-PPFRTNOVILTA---PNRSSYRVDSLLGPKNPSGIL 436
 Db 103 QLMIQEGGRLEAEALAYTKTIGEDSATHQILQTHDDPNASVLHPQALL-----A 153
 QY 437 DIDLLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPDTISLY-----DCC 486
 Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCNMPSTPSFEGIYYIEQILRHLPIS 204
 QY 487 INSILQYFQN-----NRTLLLLLTAN-----QTLMGQTSQVDWKDH 521
 Db 205 IITPLDCFWEGSQLLGESAVVIPGLNORLLWTLTNPASVMQYMKQKMSSEKISFDFTV 264
 QY 522 FLY-----C-ANAPLTFKD-----CTALALSCWADYGAPVFPF 553
 Db 265 EGYMKRAAIGSGYMEKPCNLNPNCPDTPAPKNSTQPPDVGAILSGGCGY-YAAKHMEW 323
 QY 554 ---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
 Db 324 PEELIVGGRKRNRSGLRKAQALQSVQLMTEKMYDQMDNYKVHHLGWTQEKAAEVLN 383
 QY 591 LWEEAF---LEMRAPQRMAGMFOVTTAERSLSDEINRTAEDLPIFATSYVIFLYI 647
 Db 384 AWQRNFSREVEQLLRKQSRATNYDIYVFSAAALDILAKFSPHSALSIVIGVAVTVLY- 442
 QY 648 SIALGSSYSWSRVMDYSKATLGLGVAVVLGVAVMAAGPFSVLGI----- 692
 Db 443 --APCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
 QY 693 -RSSLVI---LQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRALGRVAPSM 747
 Db 500 EQTKLILKNASTQVVPFLALGLGVDFHFI-----VGPSI 533
 QY 748 LLCISLSEAICTPLGALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKROBASRL 807
 Db 534 LFSACSTAGSPFAAAFIQVPAIKVFCLOAAIWCNSLAAALIVFPAMISLDLRRRTAGRA 593
 QY 808 DV-CCCV-----KQQLPP-----PQ 823
 Db 594 DIFCCCFPVKQKVPKVPFVPLNNGNGRGARHPKSCNNRVLPAPQNPPLLBQRADIPGS 653
 QY 824 GEGL-----LLGFFQKAYAFPLHMTITRGVWLLFLALPGVLSYMSCHTSVGLDQSLALPK 879
 Db 654 SHSLASPSLATFAFQHYTFPLMRSMVKVLTWVGFLAALISSLYASTRLQDGLDIDLVPK 713
 QY 880 DSYLLDYFLNRPVEVGAPVYFVTTLGYNFSSEAGMNAICSSACGNPFSFTQKIQYATE 939
 Db 714 DSNHKKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQQQLLRD 752
 QY 940 FPEQSYLAIP-----ASSWDDIDIMLTPSSCCCLYISGPNKDKFCFSTVNSLNC 989
 Db 753 YHDS--FRVPHVKNNDNGGLPDFWLLLFSEWL-----GNLQKIFDEERYDGR 798
 QY 990 LKNC-----MSITMGSVRPSVBO-----FHKYL-PWF 1015
 Db 799 TKECWFPNASSDAILAYKLIQVQTHVDNPNVDEKELVLTNRLVNSDGIINQRAFNYLSAWA 858
 QY 1016 LNDPRNPKCPKGLAAYSTSVNLTSDGQVLAASRFMAYHKLNKNSQD----- 1061

Db 859 TND-----VFAYG-----ASQKLPPEPRQYFHP--NEYDKIPKSLPLVYAQM 901

Qy 1062 --YTEALRAARELANITADLRKVPGTDPAPFV--FFYTITNVFEQYLITLPEGLPMLS 1117

Db 902 PFYHLGLTDSQI-KTLGHTRDLSVKYEGFGLPNYFSGIPFIEQYMT-LRSSLAMIL 959

Qy 1118 LCLVPTFAVSCLLGLDRLSGLLNLLSIVMTLVDTVGFMAWDISYNAVSLINLVSAGM 1177

Db 960 ACVLLAALVLSLLLSVMAAVLTVLSVLAIAQIFGAMTLLGKLSAIPAVILILSVGM 1019

Qy 1178 SVEFVSHITRSFALSTKFTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQI 1237

Db 1020 MLCF--NVLLISGFMTSGVNRQRRVQLSMQSLGSLVHGMLTSGVAFMLSTSPFVIR 1077

Qy 1238 FFFRLNLLITLGLLHGLVFLVFLVILSYVGPDPNFPALALEQKRAEBAVAAVMVASCNPHS 1297

Db 1078 HFCWLLLVLCVGCAGNSLLVFPILLSMVGPEAE-LVPLE-----HPD 1118

Qy 1298 RVSTADNIVNHS 1310

Db 1119 RISTPSLPVRSS 1131

RESULT 92

AD94211

ID ADE94211 standard; protein; 1285 AA.

AC ADE94211;

XX 12-FEB-2004 (first entry)

DT Fruit fly ptc polypeptide.

DE Fruit fly; patched gene; ptc; developmental abnormality;

XX Proliferative disorder; tumour; oncogenic patched mutation;

KW Patched-associated phenotype; basal cell nevus syndrome; BCNS; carcinoma;

KW meningioma; fibroma; cancer.

XX Drosophila melanogaster.

OS US6551782-B1.

XX 22-APR-2003.

XX 28-NOV-2000; 2000US-00724631.

XX 07-OCT-1994; 94US-00319745.

PR 06-OCT-1995; 95US-00540406.

PR 31-MAY-1996; 96US-00658055.

PR 22-AUG-1997; 97US-00918658.

XX (STRD) UNIV STANFORD.

PA (REGC) UNIV CALIFORNIA.

XX Scott MP, Goodrich LV, Johnson RL, Epstein E;

PI WPI; 2003-615308/58.

XX Assay for phenotyping patched status of cell in mammalian (preferably human) cell sample, involves detecting presence or absence of aberrant modification or mutation of a patched gene, and mis-expression of patched gene.

XX Disclosure; SEQ ID NO 6; 57pp; English.

XX The invention relates to an assay for phenotyping the patched status of a cell, involving detecting in vitro, in a sample of mammalian cells, the presence or absence of a genetic lesion in a patched gene (ptc) characterised by at least one of an aberrant modification or mutation of a patched gene or mis-expression of the patched gene. The invention also relates to a method for diagnosing a genetic predisposition of an animal to a developmental abnormality or a proliferative disorder, where the developmental abnormality or proliferative disorder correlates with

CC aberrant expression or activity of a patched gene or gene product,

CC involving detecting in vitro the presence of a predisposing mutation in a patched gene in cells of the animal. Characterising the phenotype of a tumour, involves detecting the presence of an oncogenic patched mutation in cells of the tumour, where the presence of the oncogenic mutation indicates that the tumour has a patched-associated phenotype. The assay is useful for phenotyping the patched status of the cell in a mammalian cell sample obtained from a human patient. The disorders that can be detected include basal cell nevus syndrome (BCNS), carcinomas, meningiomas and fibromas. This sequence represents a fruit fly ptc polypeptide of the invention.

XX

SQ Sequence 1285 AA;

Query Match 6.3%; Score 438.5; DB 7; Length 1285;

Best Local Similarity 20.0%; Pred. No. 1.8e-30;

Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

Qy 321 KGTSLSDKLSFSTHTLIGQFFQGWGTWASWPTILVLSVIPVVAAGLVFTLTDPV 380

Db 43 KGKAGSRRTAIIYLSVFSQSHLETGSSVQKHGKVLFAVLVSTFCVGLKSAQIHSKVH 102

Qy 381 ELMSAPNSQARSEKAFHQHFG-PPFTNQVILTA---PNRSSVRYDSLLILGPKNFSGIL 436

Db 103 QLWIQEGGRLEALAYTKQTIGEDSATHQLLIQTTHDPNASHVLPQALL-----A 153

Qy 437 DLDLLLELLELQERLRLHQLVQWSPQAQRNLSLDICVAPLNPNTSLY-----DCC 486

Db 154 HLEVLRKATAVKHLYDTE-W-----GLRDCNMFSTPSPGIYVIEQILRHLPIS 204

Qy 487 INSLLOYFQN-----NRTLLLTAN-----QTLMGQTSQVDKDH 521

Db 205 IITPLDCFWEGSQLGPESAVVIGLNLQRLMTLTPASVMQYKQKSEKIFDFETV 264

Qy 522 FLY-----C-----ANAPLTFK-----GTALALSCMADYCAPVPPF 553

Db 265 EQYMKRAAIGSYMEKPLNPNCPDTPAKNKSSTOPPDVGAILSGCYG-YAAKEMHW 323

Qy 554 ---LAIGGYK--GKDYSEALIMTFSL-----NNYPAGDPLRAQAK----- 590

Db 324 PEELIVGGRKRNRSGLHRLKQAQLSVVQLMTEKEMYDQWQDNKYVHLGWTQKAAEVLN 383

Qy 591 LWEEAF---LEEMRAFQRRMAGMFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYI 647

Db 384 AWQRNFSREVQLLRKQSRATNYDIYVFFSAALDDILAKFSHPALSIVIGVAVTVLY- 442

Qy 648 SLALGSYSSSRVMDSKATLGLGVAVVLGVAAMGFFSYLGI----- 692

Db 443 --AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAGLGLSALLGLVFNALTAAYAESNR 499

Qy 693 -RSSLVI---LQVVPFLVLSVGADNIFIVLEYQRLPRRFBEPREHVHGRALGVAPSM 747

Db 500 EQTKLILKNASTQVVPFLALGLGVDFHFI-----VGPSI 533

Qy 748 LLCSLSEACFPFLGALTMPAVRTTALTSGLAVILDFLQMSAFVALLSLDSKQOEASRL 807

Db 534 LFSACSTAGSFFAAAFIPVPAIKVFCLOAAIWMCSNLAAALVFPAMISLDLRRRTAGRA 593

Qy 808 DV-CCCV-----KQOELPP-----PGQ 823

Db 594 DIFCCCFPVKQKQKVPVPLPLNNGNGRAGHPKSCNNRVPLPAQNPLLRQADIPGS 653

Qy 824 GEGE-----LLGFFQKAYAPFLHWHITRGVVLVLLFALFGVLSYMSCHISVGLDQELALPK 879

Db 654 SHSLASPSLATPAFQHYTPFLMRSWVKFLTVMGFLAALISLASYASTRLQDGLDIDLVPK 713

Qy 880 DSYLLDYFLFLNRYFEVGPVYFVTTLGYNFSSSEAGMNAICSSAGCANNFSFTKIQVATE 939

Db 714 DSNEHKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQQQLLRD 752

Qy 940 FPEQSYLAIP-----ASSWDDDFIDMLTPSSCCRLYISGPNKFCFSTVNSLNC 989

Db 753 YHDS---FRVPHVINKDNGGLPFWLLLFSEWL-----GNLQKIFDEYRDGRL 798

QY 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TRECFNPNASSDAILAYKLIQVGHVNDVDEKELVLTNRLVNSDGLINQRAFYVLSAWA 858
QY 1016 LNDNRNPKCPKGLAAYSTSVNLTSDQVQLASRFMAYHKPLKNSQD-----1061
Db 859 TND-----VFAYG-----ASQGLYPEPRQYFHP--NEYDLKIPKSLPLVYAQM 901
QY 1062 --YTEALRAARELAANITADLRKPGCTDPAFV--FPYITVNFVEQYLTILPEGLFMLS 1117
Db 902 PFLHGLTDTQI-KTLIGHIRDLSVKYEGFGLPNYPSGIPFWEQWT-LRSSLAMIL 959
QY 1118 LCLVPTFAVSCLLGLDRSGLLNLSIVMILVDTVGFMALWDISYNAVSLINLSAVGM 1177
Db 960 ACVLLAALVLSLLSVMAAVLIVLSLASIAQIFGAMTLGKLSAIPAVILISVGM 1019
QY 1178 SVEFVSHITRSFAISKPTWLRKAETISGSAFVAGVAMTNLPGILVLGLAKAQLQI 1237
Db 1020 MLCF--NVLLISLGFMTSVGNRRQVQLSMQMSLGLPLVHGMLTSGVAVFMLSTSPFEFVR 1077
QY 1238 PFRNLNLTLLGLHGLVFLPVILSYGPDVNPALALEOKRABEABAAMVAVSCPNHPS 1297
Db 1078 HECWLLVVLGVACNSLLVFFILLSMVGPEAE-IVPLE-----HPD 1118
QY 1298 RVSTADNIYVNH 1310
Db 1119 RISTPSPVRS 1131
RESULT 93
ADH62718
ID ADH62718 standard; protein; 1285 AA.
XX AC ADH62718;
XX DT 25-MAR-2004 (first entry)
XX DE Fly patched (PTC) homologue protein.
XX KW PTC; patched; embryonic development; cellular regulation;
XX KW signal transduction; fly.
XX OS Drosophila melanogaster.
XX PN US2003148388-A1.
XX PD 07-AUG-2003.
XX PF 03-JAN-2001; 2001US-00754032.
XX PR 07-OCT-1994; 9AUS-00319745.
XX PR 06-OCT-1995; 95US-00540406.
XX PR 20-OCT-1997; 97US-00954669.
XX PA (STRD) UNIV STANFORD.
XX PI Scott MP, Goodrich LV, Johnson RL;
XX DR WPI; 2003-897566/82.
XX PT New DNA sequences encoding patched protein from organisms including human
XX PT and mouse are used to elucidate embryonic development and cellular
XX PT regulation associated with patched signal transduction and identify
XX PT patched ligands.
XX PS Disclosure; SEQ ID NO 6; 40pp; English.
XX CC The present invention relates to a DNA sequence encoding a patched (PTC)
XX CC gene. The invention is used to elucidate embryonic development, cellular
XX CC regulation associated with signal transduction by the patched gene,
XX CC identify agonists and antagonists of that signal transduction, identify
XX CC and isolate ligands of patched protein and assay for levels of

CC transcription of the patched gene. The present sequence is fly patched
CC protein.
XX
SQ Sequence 1285 AA;
Query Match 6.3%; Score 438.5; DB 7; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.8e-30;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;
QY 321 KGTSLSDKLSFSTHTLLGQFFQGCWCTWASWPLTILVLSVIPVVALAAGLVFTELTDPV 380
Db 43 KKGARSIAIYLRVSFQSHLETGSSVQKRGKGVFLVAILVLSFCTCVGLKSAQIHSKVH 102
QY 381 EIMSAIPNSQARSEKAFHDQHFQ-PPFRTNQVILTA---PNRSSYRYDSILLGPKNFSGL 436
Db 103 QLMIQEGGRLEAELAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLEQLERLHQVMSPEAQRNISLQDICVAPLNPDNTSLY-----DCC 486
Db 154 HLEVILKATAVKVHLDTSE-W-----GLRDMCMFSTPSEFGIYIEQILRLHLP 204
QY 487 INSLIQYFQN-----NRTLLILLTAN-----OTLMGQTSQVDMKH 521
Db 205 IITPLDCFWESQLLGPESAVVIPGLNQLRLTTLNPNASVMQMKQSEKISDFETV 264
QY 522 FLY-----C-----ANAPLTFKD-----GTALALSMDYDGPVFPF 553
Db 265 EQYMKRAAIGSYMEKPCINPLNCPDTPAKNSTQPPDVGAILLGGCYG-YAAKHMHW 323
QY 554 ---LAIGYK---GNDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----590
Db 324 PEELIVGGRKRNRSGLRKAQALQSVQLMTEKEMYDQMDNKKVHHLGWTQEKAAEVLN 383
QY 591 LWEEAP---LSEMRAPORMAGMFOVTTAERSLSEDEINRTTAEDLPATSVIYFLVI 647
Db 384 AWQRFSEVSQLLRKQSRITATNYDIYVFSSAALDDILAKSHPSALSIVIGAVTVLY- 442
QY 648 SIALGSYSSWRVMVDSKATILGGLGVAVVLGAVMAAMGFFSYLGI-----692
Db 443 --AFCTLLRW-RDPVRGQSSGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI---LQVVPFLVLSVGADNIFIVLEYQRLPRRPGEPREVHIGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVPFLALGLGVDFIFI-----VGPSI 533
QY 748 LLCSLSEACFELGALTMPBAVTRTALTSGLAVILDFLLQMSAFVALLSLDSKROEASRL 807
Db 534 LFSACSTAGSFFAAAFIVPALKVFCLQAAIVMCSNLAAALLVFPAMISLDLRRRTAGRA 593
QY 808 DV-CCCV---KQOELPP-----PGQ 823
Db 594 DIFCCCFPVWKEQPKVAPVPLPLNNGNGRGARHPKSCNNRVPLPAQNPLLRADIPGS 653
QY 824 GEGE-----LLGFFQKAYAPFLHMTIRGVLLFLALFVGLSYSMCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFOHYTPFLMRSMVKELTVMGFLAALISSLYASTAQDGLDIDLVK 713
QY 880 DSYLLDYFLPLNRYPEVGAPVYFVTLGLNPSSEAGMAICSSACNNEFSTQKQIYATE 939
Db 714 DSNHFKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQOQLLRD 752
QY 940 PPEQSYLAIP-----ASSWVDDFDIMLTPSSCCRLYISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNDNGGLPDFWLLLFSEWL-----GNLQKIPDEEYRDGEL 798
QY 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLIQVGHVNDVDEKELVLTNRLVNSDGLINQRAFYVLSAWA 858
QY 1016 LNDNRNPKCPKGLAAYSTSVNLTSDQVQLASRFMAYHKPLKNSQD-----1061
Db 859 TND-----VFAYG-----ASQGLYPEPRQYFHP--NEYDLKIPKSLPLVYAQM 901

QY 1062 --YTEALRAARELAANITADLRKVPGTDPAPFV--FPYTTNNVYEQYLITLPEGLPMLS 1117
Db 902 PFYHLGLTDTSGI-KTLIGHIRDLDSVKYEGFGLPNYPSGIPFIWEQYMT-LRSSLAMIL 959
QY 1118 LCLVPTFAVSCLLGLCLDRSGLNLLSTVMTLVGFMALWDISYNVAVSLINLVAVGM 1177
Db 960 ACVLAALVLSLSSVMAAVLVLSVLSLAQIFGNTLGLKLSAIPAVILSVGM 1019
QY 1178 SVEFVSHITRSPATSKPTWLERAKEATISMGSAVFAVAMTNLPGILVLGLAKAQIUI 1237
Db 1020 MLCF--NVLSISGFMTSVGNRRVQLSMQSLGSLPLVHGMLTSGVAVFMLSTSPFVIR 1077
QY 1238 PFRRLNLLITLGLLHGLVFLPVILSVYGPDPNPALEQKRAEAVAVMVASCPNHP 1297
Db 1078 HFCWLLLVLCVACNSLLVFPILLSMVGPEAE-LVPLE-----HPD 1118
QY 1298 RVSTADNIYVNH 1310
Db 1119 RISTSPFVRS 1131

RESULT 94
ID ADE48976 standard; protein; 1285 AA.

AC ADE48976;
XX
XX
XX 29-JAN-2004 (first entry)
XX
XX Drosophila patched protein (ptc).
XX
XX cytostatic; vulnery; gene therapy; phenotyping; patched status;
KW patched gene; genetic predisposition; basal cell nevus syndrome; tumour;
KW carcinoma; meningioma; medulloma; fibroma; cancer; wound healing; aging;
KW fruit fly; patched gene; ptc.
XX
OS Drosophila melanogaster.
XX
XX US2003186309-A1.
XX
XX
XX 02-OCT-2003.
XX
XX 22-APR-2003; 2003US-00421446.
XX
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX 31-MAY-1996; 96US-00656055.
XX 22-AUG-1997; 97US-00918658.
XX 28-NOV-2000; 2000US-00724631.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX (REGC) UNIV CALIFORNIA.
XX
XX Scott MP, Goodrich LV, Johnson RL, Epstein E;
XX
XX WPI; 2004-041193/04.
XX N-PSDB; ADE48975.
XX
XX Phenotyping the patched status of a cell for diagnosing a genetic
XX predisposition for a tumor comprises detecting the presence or absence of
XX a genetic lesion having aberrant modification, mutation or mis-expression
XX of the patched gene.
XX
XX Disclosure; SEQ ID NO 6; 60pp; English.
XX
XX The invention describes an assay for phenotyping the patched status of a
XX cell comprising detecting in a sample of mammalian cells the presence or
XX absence of a genetic lesion having aberrant modification or mutation of a
XX patched gene or mis-expression of the patched gene. The assay is useful
XX for diagnosing a genetic predisposition of an animal, e.g. basal cell
XX nevus syndrome, predisposition for developing tumour, i.e. carcinoma,
XX meningioma, medulloma, or fibroma. A genetic construct encoding a patched
XX polypeptide is used to treat an animal having a disorder comprising loss

CC of function of a wild-type patched gene, such as cancer, and can enhance
CC patch function in e.g. wound healing and aging. This sequence is encoded
CC by the fruit fly patched gene.
XX
SQ Sequence 1285 AA;
Query Match 6.3%; Score 438.5; DB 8; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.8e-30;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;
QY 321 KGTSLDKLSFSTHTLLGQFQGWGTWASWPTILVLSVIFVVALAAGLVFTELITDPV 380
Db 43 KKGARSRTAIYLSVFSQSHLETGLSSVQKAGKLVFVAILVLSVTCFCVQKSAQIHSKVH 102
QY 381 ELMSAPNSQARSKAFHQHFG-PPFRTNOVILTA---PNRSSVRYDLSLLGPKNFGSIL 436
Db 103 QLMIQEGGRLEAELAYTKTIGEDSATHQLLQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLLELLELQBLRHLQVMSPEAQRNISLODICYAPLNPDNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHYDTE-W-----GLRMCNMPSTSPGIIYIIEQLRHLIPCS 204
QY 487 INSLLYQFQN-----NRTLLILLTAN-----QTLMGQTSQVDMKDH 521
Db 205 IITPLDCFWEGSOLLGPESAVVIFGLNORLLTTLNPASVMQYMKQKSEKISFDFTV 264
QY 522 FLY-----C-----ANAPLTFKD-----GTALALSCMADYGAPEPF 553
Db 265 EQYMKRAAIGSYWEKPCPLNPNPCDTPAPKNKSTOPPDVGAILSGCYG-YAAKMHW 323
QY 554 ---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAQ----- 590
Db 324 PEELIVGRKENRSGHLRKAQALQSVVQLMTEKEMYQWQDNVYVHHLGWTQEKAAEVLN 383
QY 591 LWEEAF---LEEMRAFQRMAGMEQVITFAERSLEDEINRTAEDLPFATSIYIVILYI 647
Db 384 AWQNFREVEQLLRKQSRITATNYDIYVFSSAALDDILAKFSPSALSIVIGVAVTDLY- 442
QY 648 SLALGSYSSWSRVMVDKATLGLGVAVVLGAVMAAMGFYSYLG----- 692
Db 443 --AFCTLLRW-RDPVGGSSVGAVGLLMCFSTAAGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI---LQVVPFLVLSVADNIPFIVLEYQRLPRRCEPREVHIGALGRVASM 747
Db 500 EQTKLILKNASTQVVPFLALGLVDHIFI-----VGPSI 533
QY 748 LLCSLSEACFELGALTMPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKROEASRL 807
Db 534 LFSACSTAGSFFAAFIPVPALKVFCLOQAIVMCSNLAALLVPAMISLDRRTAGRA 593
QY 808 DV-CCCV-----KPOELPP-----PGQ 823
Db 594 DIFCCCFPWKEQPKVAPVPLPLNNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGS 653
QY 824 GEGE-----LLGPFQKAYAPFLLHWTIRGVVLLLFALFGVLSYMSCHISVGLDQELALPK 879
Db 654 SHSLASFSLATPAFQHYTPFLMRSWVKFLTVMGFLAALISLYASTRLQDGLDIDLVPK 713
QY 880 DSYLDYDFLNLRYPEVGAPVYVFTTLGVNFSSEAGMNAICSSAGCNPFSTFKIQYATE 939
Db 714 DSEHKFLDAQTRLF--GFYSMAVTOG-----NFEYPTQQQLLRD 752
QY 940 FPEQSYLAIP-----ASSWYDDFDIMLTPSSCCRLYISGNKOKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVTKNDNGGLPDLFLLFSEWL-----GNLQKIFDEEYRDSGL 798
QY 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAIAYKLIQVQTHVDNPDVKELVTLNRLVNSDGLIINORAFNYLSANA 858
QY 1016 LNDPRNPKCPKGLAAYSVNLTSQGVLSRSMFAYHFKPKNSQD----- 1061
Db 859 TND-----VFAYG-----ASQGLYPPEPRQYFHQP--NEYDKLPKSLPLVYAQM 901

QY 1062 --YTEALRAARELANITADLRKVPCTDPAFV--FPYITITNVFEQYLTILPEGLFMLS 1117
 Db 902" PFYTLGLTDSQI-KTLIGHIRDLSVKYEGFLENYPGIPFIFWEQYMT-LRSSLAMTL 959
 QY 1118 LCLVPTFAVSCLLGLDLSGLNLLSYWILVDTVGFMAWDISYNAVSLINLSVAGM 1177
 Db 960 ACVLLAALVLSLLSVAWVILSVLASAQIFGAMTLGKLSAIPAVILLSVGM 1019
 QY 1178 SVEFVSHIRSFSAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQI 1237
 Db 1020 MLCF--NVLLISGFMTSGNRQRRVQLSMQMSLGPLVHGMLTSGVAVFMLSPPFVIR 1077
 QY 1238 PFRNLNLTLLGLLHGLVFLPVLSYGVDPVNPALALEQKRAEAAVAVMVASCPNHS 1297
 Db 1078 HFCMLLVLCVGCACNSLLVFPILLSMVGPEAB-LVPLE-----HPD 1118
 QY 1298 RUSTADNIVNHS 1310
 Db 1119 RISTPSPLEVRSS 1131

RESULT 95

ADN22757

ID ADN22757 standard; protein; 714 AA.

AC

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 714 AA;

Query Match 5.5%; Score 379; DB 8; Length 714;
 Best Local Similarity 23.0%; Pred No. 2.5e-25;
 Matches 173; Conservative 119; Mismatches 299; Indels 160; Gaps 25;
 QY 545 DYGAFFVFPFLAIGGYKGDYSEAEALIMTFSLN--NYP-----AGDPRLAQ--AKLWEEAF 596
 Db 58 DYAAEM-----TGCGSG-----FASNVLNWPEDMILGNPRAKKGKLSGADA 100
 QY 597 LEEMRAFORRMAGMQVTFRTAERSLEDEINRTAEDLPFATSYIVIFLY-----646
 Db 101 LQ-----FVFLVASPADVFLRFKQSLSEB-----TLETLSYERHCILHHRHPLILCSNSV 151
 QY 647 ---ISLALGSVSSMSRVWV-----DSKATLGLGVAVVILGAVMAAGFFSYL 690
 Db 152 NSTPSSLLDMLSCHWMLLLILDUTAFRLQHWGFGALAGLVVTFASVAGIGLATWF 211
 QY 691 GIRSLVILQVVPFLVLSVGADNIFIVLEYQLRRPPEPPEPPEVHIGALGRVAPSMILJC 750
 Db 212 GLEFNAATTQIVPFLTIGVDNMPMLLHNYRDVVKLAGGHAEMAI--LMRETGHSILCT 269
 QY 751 SLSEAI CFFLGALTMPAVRTFALTSLGLAVILDFLQMSAFVALLSLSKQREARLDVC 810
 Db 270 SINNLISFLTGTLLPILALRSFCQSSILLTFNFIALTITPAIISIDLRKKAKORRDF 329
 QY 811 CCVK--PQELPPPGQEGE-----LLGFFQKAVAPFLHWHITRGVLLLLFL 854
 Db 330 TREESTSEASIMQDGTQOMASSDDPAPWSLHSFIRYIYPIPSKPAKVAIVGCC 389
 QY 855 ALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGPVYFVTLGYNFS--- 911
 Db 390 ALLGASFVGMRSQSTLGLSLGDLVPEHTAPAQFLRADKYFSP-YPM-FAVIKGPNIIDYAH 447
 QY 912 SEAGNNAICSSAGCNPNFPTOKIQVATEPPQSYLAIPASSWDDFIDWLTPSSCCRLYI 971
 Db 448 QQRQIDNTRQSIGSSKYVKNK---NEEPSEKY-----WGLMRDWLI--SIQRGFD 494
 QY 972 SGPNKDKP-----CPSTVNSLNCCKNCSMTGMSVRPSVEQFHK--YLPWFNDRPNI 1022
 Db 495 BEVAKGSFNLSTGTVIGNSVEDARLAHALMCHSGSLFECAGRIKIRLIPPLYT----- 549
 QY 1023 KCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRK 1082
 Db 550 -----GLTDTAVIVDAIKDIRSVCEF-----571
 QY 1083 VPGTDPAPFVPPYITINVFYQYLTILPEGLFMLSCLVPTFVAVSCLLLGLDLRLGLNL 1142
 Db 572 ---TDQGLPNFPQGIATFWEQYFLTCNLMAQISIIITISVFCVSVLL-FNPMAALMVV 627
 QY 1143 LSIWMLVDTVGFMAWDISYNAVSLINLSVAGMSVEFVSHITSPASTKPTWLERAK 1202
 Db 628 CILGIMTCELAFGMGLVGKILNPVSAVTLITAVGIGVEFTVHVVSFL-----TALGTRS 682
 QY 1203 BATISMGSAVPAV---AMTNLPGILVLGLA 1230

New recombinant DNA construct comprising a promoter positioned to provide
 for expression of a polynucleotide encoding a polypeptide from a
 microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 5410; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
 promoter functional in a plant cell, where the promoter is positioned to
 provide for expression of a polynucleotide encoding a polypeptide from a
 microbial source. The invention also relates to a transformed plant
 comprising the recombinant DNA construct and a method of producing a
 transformed plant having an improved property. The plant is a crop plant
 such as maize or soybean. The method of producing a transformed plant

Modified-site 192..195
/note="N-glycosylation site"
Modified-site 194..196
/note="Protein kinase C phosphorylation site"
Modified-site 200..202
/note="Protein kinase C phosphorylation site"
Modified-site 201..204
/note="cAMP and cGMP-dependent protein kinase phosphorylation site"
Modified-site 215..218
/note="Casein kinase II phosphorylation site"
Modified-site 219..222
/note="Casein kinase II phosphorylation site"
Modified-site 225..228
/note="Casein kinase II phosphorylation site"
Modified-site 230..233
/note="Casein kinase II phosphorylation site"
Modified-site 275..278
/note="N-glycosylation site"
Modified-site 279..282
/note="N-glycosylation site"
Modified-site 307..312
/note="N-myristoylation site"
Modified-site 329..335
/note="Tyrosine kinase phosphorylation site"
Domain 383..540
/note="Sterol-sensing domain, SSD, domain"
Modified-site 418..422
/note="N-myristoylation site"
Modified-site 504..509
/note="N-myristoylation site"
Modified-site 508..510
/note="Protein kinase C phosphorylation site"
Modified-site 530..533
/note="N-glycosylation site"
Modified-site 535..540
/note="N-myristoylation site"
Modified-site 541..544
/note="Amidation site"
Modified-site 561..563
/note="Protein kinase C phosphorylation site"
Modified-site 572..575
/note="Casein kinase II phosphorylation site"
Modified-site 597..600
/note="Casein kinase II phosphorylation site"
Modified-site 662..664
/note="Protein kinase C phosphorylation site"
Modified-site 678..681
/note="N-glycosylation site"
Modified-site 681..688
/note="Tyrosine kinase phosphorylation site"
Modified-site 692..695
/note="N-glycosylation site"
Modified-site 737..740
/note="N-glycosylation site"
Modified-site 740..743
/note="Casein kinase II phosphorylation site"
Modified-site 746..748
/note="Protein kinase C phosphorylation site"
Modified-site 759..761
/note="Protein kinase C phosphorylation site"
EP1229046-A2.
07-AUG-2002.
28-JAN-2002; 2002BP-00001167.
30-JAN-2001; 2001WO-US000663.
30-JAN-2001; 2001WO-US000664.
30-JAN-2001; 2001WO-US000665.
30-JAN-2001; 2001WO-US000667.
30-JAN-2001; 2001WO-US000668.

PR 30-JAN-2001; 2001WO-US000669.
PR 23-MAY-2001; 2001US-00864761.
PR 09-OCT-2001; 2001US-0327898P.
PA (AEOM-) AEOMICA INC.
XX Zhan J;
PT
XX WPI; 2002-676582/73.
DR N-PSDB; ABV78761, ABV78762.
XX
PT Novel isolated human testis expressed Patched like protein (HTPL), useful
for identifying agonist and antagonist and specific binding partners, and
for treating subjects having defects in HTPL.
XX
PS Claim 13; Fig 4; 718pp; English.
XX
CC The present sequence is the protein sequence for human testis expressed
Patched like protein (HTPL) short form (HTPL-S). HTPL has two isoforms,
with a few single base pair differences between the two. One of the
CC single base pair changes introduces a premature stop codon in HTPL-S (S
for short) compared to HTPL-L (L for long). HTPL shares an overall
CC structure organisation with the Patched protein. The shared structural
CC features strongly imply that HTPL plays a role similar to that of
CC Patched, and is a potential tumour suppressor. HTPL is important in
CC regulating male germ cell development, and the HTPL gene was mapped to
CC human chromosome 10p12.1. HTPL and its coding sequence are useful for
CC diagnosing a disorder caused by mutation in HTPL, and in therapy and
CC manufacture of a medicament for treatment or prevention of such disorder
CC associated with decreased expression or activity of human HTPL. Such
CC disorders include disorders of testis, or adrenal, adult and foetal
CC liver, bone marrow, brain, kidney, lung, placenta, prostate, skeletal
CC muscle or colon function. HTPL proteins and nucleic acids are clinically
CC useful diagnostic markers and potential therapeutic agents for male
CC infertility and cancer
XX
SQ Sequence 767 AA;
Query Match 5.1%; Score 351.5; DB 5; Length 767;
Best Local Similarity 20.5%; Pred. No. 1.1e-22;
Matches 158; Conservative 137; Mismatches 285; Indels 189; Gaps 31;
QY 334 HT-----LLGQFFQGGTGWASWPLTILVLSVIPVVAALAGLVF--TELTTPDVELWSAP 386
DB 115 HTDCLGLLSRTFFQMLGQVGAHPWIFLLAPLMLTAALGTGFLYLPKDEEDLEHYTPV 174
QY 387 NSQARSEKAFHDQHFQFPFRITNQVILTAPNRSS-----YRDSLLLPKPNPSGIL 436
DB 175 GSPAKAERRFVQGH---FTTNDYRFSASRRSTEANFVSLVSVSYSDSLDPATPAEVS 230
QY 437 DLDLLELLELQERLRLHLQVMSPEAQRNLSLODIC--YAPLNPNTSLYDCCI--NSLLQ 492
DB 231 KLDGAVQDLRV-----AREKSGIQYQVQVARYAL-----CVPPPIIY 270
QY 493 YFQNNRTLLLTANOTLMGQTSQVDKHFYCANAPLTFKDGTTALALSMDADYGA-PVF 551
DB 271 AMQVKNLTNL-----SSISPPAY-----NHGRHPLY 296
QY 552 PFLATGGY-----KGKDYSEABALIMTFSLNNYPAGDPRL-AQAKLWEPAFLBEMRAP 603
DB 297 LTGPFPGYILGSLGMGQILLRAKMRLLYLLK---TEDPEYDVQSQKQLTHLLDQFTMI 353
QY 604 QRRMA--GMFQVTFTA-BRSLEDEINRTTAEDLPPIPATSVYIVFLYISALSGSYSSWSRV 660
DB 354 KNILAKKIEVHVHTSLRQLEFEATSVTV--IPVFLAYILILILFVNTSCFRFD----- 406
QY 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGRSSILVILQVVPFVLVSGADNIFTFVLE 720
DB 407 CIRNKMCAVAFGVISAFVAVSGFGLLHIGV-PFVLIIVANSFPLILGVVDMMFIMISA 465
QY 721 YORLPFRPGEPREVIHGRALGRVAVSMLLCSUSEAICPFLGALTMPAVRTFALTSGLAV 780
DB 466 WHK-TNLADDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSQVCCTVTGMTL 523

CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
CC protein.

AA		Sequence 783 AA;	
SQ			
	Query Match	5.1%; Score 349; DB 7; Length 783;	
	Best Local Similarity	21.5%; Pred. No. 1.9e-22;	
	Matches	193; Conservative 160; Mismatches 335; Indels 208; Gaps 42;	
Qy	444	LLEQLERLRLHLYWSPEAQRNISLDICVAPLNPDNTSLVDCCINSLLQVFON---NRT 499	
Dd	10	ILKHAATVKIQV--PRPGNYTFAHICI--LNDKTCIADDIVH-VLELKVARATNRT 64	
Qy	500	LLLILANTQTLMGOTSQVDWKDHFPLYCANAPLT-FKDGTALALSCMADYGAPVPFFLAIGG 558	
Dd	65	NPAIT-----YPIHLKDGRA-----VTNGHQLGCG 89	
Qy	559	--YKGKD-YGEAELIMTFSLANNYPAGDPRLQAQAKWESEAFLEEMRAPORMAGMQVTP 615	
Dd	90	VTVHSKDRVSASAEAIQITYYLOGINSLNDMAVER--WESSFCDTVRLFQKSNSKVMPY 147	
Qy	616	TAEBSLEDEINRTTAEDLPIFATSYIVIFYLSIALGSYSMSRWMDSKATLGLGVAV 675	
Dd	148	TSS-SLRDDPKGRVSERYLVTSLLV---VTMAILCS--MODCVSRKPWLGLLGLVT 201	
Qy	676	VLGVAMAAMGFYSIGIRSSLVLQVVPFLVLSVGADNIIFVLEYORLPRRCEPREVH 735	
Dd	202	ISLATLTAAGIINLTGGKYNSTFLG-VPFVMGLHGXYGTEFMLS--RKTRDQHVK 256	
Qy	736	IGRALGRVASMLLCSLSEA---ICPFLGA-LTPMPAVRTEALTSGVALILOFLQMSA 790	
Dd	257	-ERTAAYVADMLSFSLTWTMYLVTFGIGASPPTIEAARPCNSCIAIPFNLYLVSP 315	
Qy	791	FVALISLDSKRQEASRLDVCCVKP----QELP-----PPQOGE----- 825	
Dd	316	YGSSLVFTGYENNYQHSIFCRKVPKEALQEPARYRFLTARTFSEDTAEGEANTYES 375	
Qy	826	GLLLGFPQKAYAPFLHWITRGVVLLPLALFGVLSYSMCHISVGLDQELALPKDSYLLD 885	
Dd	376	HLLVCFLKRYXCWITTYVKPFVLFYLIYISPALMGYLQVSGSOLSNIVATOTIE 435	
Qy	886	YFLFLNRVFEVGPVYFVTTLTGYNFSESAGNAICSSAGCNCNFSTOKIOVATEFPBQS 945	
Dd	436	YTTHQQKFYSNPVI-----GFIYESIEYWNYSVQEDV 470	
Qy	946	LAIPAS-----SWVDFFIDMLTPSSCCRLLYIS-GPNKKDKFCPSTVNSLCLKNCHSIWTGS 1000	
Dd	471	LEYTKGFVRISWFESYLNLYR-----KLNVSTGLPKKNFTDLNRSP--LK----- 514	
Qy	1001	VRPSVEQHKYLPWFNLDRPNIKCPKGLAAYSTSVNLTDGQVLASREMAHYHKLK-NS 1059	
Dd	515	-----APOSFHFQEDIIFSK-----KYNDEVD-----VASMFLVAKTMETNR 553	
Qy	1060	QDYTEALARARELANIATADRLKPGTDPAPEVFPYTIITNVFYEQYLITLPEGFLMGLSC 1119	
Dd	554	EELYDLLLETURL--SVTSKVKFIV-FNPSF-----VYMDRVASSL-GAPLHNCS 599	
Qy	1120	LVPTFAV--SCLLGLDLRSGLNLN---LSIVMLIVDTVGMALWDISYNAVSLINLYSA 1174	
Dd	600	ISALFULLFESAFLVA---DSLINVITLVTVSVBEFVIGFMTLMKVELDCISVLCCLIYG 655	
Qy	1175	VGMSVBEFVSHITRSPAIS---TKPTWLERAKEATISMGSVAFGAVAMTN-----LPGILV 1224	
Dd	656	INYTTIDNCAPMLSTFVLGKDFTRTKWRNALLEV-----HGVALIOSLYCIYVGLIP 706	
Qy	1227	LGLAKAQITQIFFFRNLNLIITLGLLHGLVFLPVILSYVGDVNPALAEQGRAEE 1282	

```

781 ILDFLLQMSAFVALLSLDSKRQEQASRLDYCCCVKQELP-----PPGOG 824
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
524 LFCYFNITCFGAFMALDGREV-----VCLCWLKAKDKWSPKPCFPFGSVDEHG 578
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
825 EGU--LLGPFQKAYAPPLLHWIRGVLLFLFLAFGLVGLSMCHISVGLDQELALPKDSY 882
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
579 TDHPTSLFRDFVGFPLTSESKYFVVFVLYIISIIYGCYFHVQEGDLRLNLASDDSY 638
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
883 LLDYFLNLNYP-EVGAPVYFVTTLGVNPFSSSEAGMNAICSSAGCNMFSFTQIKQYATEFP 941
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
639 ITPFVFNVEENYFSDYGRVMVIVTKKYDYWDK-----DVRQKLECNCTKIF 683
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
942 EQSVLAIP--ASSWDDFDLWLTSPSSCCRLYISGPNKDKFCPSVTNSLNCNCKMCSITMG 999
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
684 EKNVYDKNUTEFWLDIAYQ-----YLGKNSD--PNEKNT-----717
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1000 SVRPSVEQFHKLFPWLFLNDRPNIKCPKGLAAAYSTSVNLTSDDGVLASR 1048
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
718 -----FNNNI PDEFNSFPN-----FOHDINISSNBSNEISSR 748
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 98	
ADI211155	
ID	ADI211155 standard; protein; 783 AA.
XX	
AC	ADI211155;
XX	
DT	15-APR-2004 (first entry)
XX	
DE	Novel human protein #130.
XX	
XX	forensic; nutritional source; damaged tissue; diseased tissue;
KW	myeloid cell disorder; lymphoid cell disorder;
KW	bone cartilage tissue growth; tendon tissue growth;
KW	ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW	tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
XX	
OS	Homo sapiens.
XX	
XX	WO2003025148-A2.
FN	
XX	
PD	27-MAR-2003.
XX	
XX	19-SEP-2002; 2002WO-US029964.
PF	
XX	19-SEP-2001; 2001US-0323739P.
PR	
PR	13-SEP-2002; 2002US-00323739.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI	Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI	Haley-Vicente D;
XX	
DR	WPI; 2003-354603/33.
DR	N-PSDB; ADI21871.
XX	
XX	New polynucleotides and secreted proteins, useful for treating myeloid or
PT	lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT	tissue growth or regeneration, in wound healing, and in tissue repair and
PT	replacement.
XX	
PS	Claim 20; SEQ ID NO 406; 156pp; English.
XX	
CC	The invention relates to an isolated polynucleotide encoding a
CC	polypeptide with biological activity. The polynucleotides and
CC	polypeptides are useful in diagnostics, forensics, gene mapping,
CC	identification of mutations responsible for genetic disorders and other
CC	traits, to assess biodiversity, as nutritional sources or supplements.
CC	The polynucleotides may also be used as molecular weight markers,
CC	chromosome markers or map related gene positions, or as an antigen to
CC	raise anti-DNA antibodies or elicit immune response. The polypeptides are

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:41:11 ; Search time 53 Seconds
(without alignments)
2418.127 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAIISNPLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 80:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2385	34.5	1278	2 T30188	Niemann-Pick C dis
2	1394	20.2	1055	2 T05663	hypothetical prote
3	1347	19.5	1456	2 T15961	hypothetical prote
4	1329	19.2	1170	2 S52525	probable membrane
5	646.5	9.4	1434	2 T30172	transmembrane prot
6	628.5	9.1	1442	2 T18538	patched protein -
7	600	8.7	915	2 S44797	F09G8.4 protein -
8	598.5	8.7	933	2 T25600	hypothetical prote
9	583.5	8.4	1182	2 T13952	membrane protein p
10	573.5	8.3	1003	2 T26746	hypothetical prote
11	572	8.3	1220	2 T18291	patched protein -
12	539.5	7.8	889	2 T29590	hypothetical prote
13	514	7.4	1405	2 T27969	hypothetical prote
14	473.5	6.9	1015	2 T15830	hypothetical prote
15	464.5	6.7	1299	1 S08119	membrane protein p
16	458.5	6.6	955	2 T26112	hypothetical prote
17	436	6.3	956	2 A89153	protein C24B5.3 [i
18	409.5	5.9	936	2 T26521	hypothetical prote
19	407	5.9	800	2 T26683	hypothetical prote
20	399.5	5.8	881	2 T31739	hypothetical prote
21	381	5.5	840	2 T33217	hypothetical prote
22	379	5.5	714	2 T16126	hypothetical prote
23	367	5.3	690	2 T23399	hypothetical prote
24	367	5.3	877	2 T24097	hypothetical prote
25	349	5.1	845	2 T25657	hypothetical prote
26	347.5	5.0	820	2 T32908	hypothetical prote
27	312	4.5	890	2 T22186	hypothetical prote
28	297.5	4.3	871	2 T28706	hypothetical prote
29	294	4.3	983	2 T21213	hypothetical prote

F09G8.3 protein -
hypothetical prote
SREBP cleavage act
related to SREBP c
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
transport protein
probable integral
probable ABC-type
multidrug resistanc
hypothetical prote
multidrug efflux p
antibiotic transpo
acriflavin resista
probable efflux pu
probable multidrug
multidrug-efflux t
RND multidrug efflu
probable RND efflu
RND multidrug efflu
hypothetical prote
probable drug effl
multiple transfera
hypothetical membr
conserved hypotet
conserved hypotet
cation efflux syst
glycoprotein B pre
probable multidrug
hypothetical prote
Na+/H+ antiporter
multidrug efflux p
RND multidrug efflu
probable multidrug
acriflavin resista
cation efflux syst
conserved integral
hypothetical prote
AcRb/AcrD/AcrF fam
conserved hypotet
probable antibioti
probable RND efflu
hypothetical prote
probable membrane
conserved hypotet
ribonucleoside-dip
hypothetical prote
probable oxidoredu
probable oxidoredu
acriflavin resista
probable efflux pu
RND multidrug effl
conserved hypotet
ribulose biphosph
protein translocas
cation efflux syst
probable integral
hypothetical prote
probable RND efflu
probable RND efflu
hypothetical prote
conserved hypotet
acridine efflux pu
acridine efflux pu
acriflavin resista
CFTR protein - Afr
hypothetical 46.1K
hypothetical prote
multidrug efflux p

ALIGNMENTS

RESULT 1
T30188
Niemann-Pick C disease protein - mouse
N:Alternate names: NPC1 protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
R:Accession: T30188
R:Loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellison, Science 277, 232-235, 1997
A:Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis gene
A:Reference number: Z20765; PMID:97362324; PMID:9211850
A:Accession: T30188
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1278 <LOF>
A:Cross-references: UNIPROT:O35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:g2251241;
C:Genetics:
A:Gene: Npc1
A:Map position: 18

Query Match 34.5%; Score 2385; DB 2; Length 1278;
Best Local Similarity 40.2%; Pred. No. 1.2e-158;
Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY 14 LLLRLAQSEPYTHIQPGYCAFYDCGKXNPGLSGMLTNSVCLSNTPARKITGDHLL 73
DB 12 LLLLLCPAQVFSG-----SCWYBGC-----IATGD-----KRYNCKYSGPPKPLPKDGYDL 59

QY 74 LQKICPRLTYGTQACCSAKOLVSLASLSITKALLTRCPACSNFVNLHCHNTCSNQ 133
DB 60 VOELCPGLFF-DNVSLLCCDIQQLTKSLNQLPLQFLSKPCSCFYNLMTLFCELTCSHQ 118

QY 134 SLFINVTRVAQLGAGOLPA---VWAYEAFYQHSFAEQSDSCSRVRVPAAATLAVGTC 189
DB 119 SQFLNVTATEDYFDKTPENKTNVKELEYVYQGSFANAMYNACRDVEAPSSNEKALGLLC 178

QY 190 GYVGSALCNAQRLNFGQDTGNGLAPLDI-----TFHLEPGQAVGSGIQPLNEGVARN 244
DB 179 GRDARA-CNATNWTYMFKNKGQAPFTIIPVSDLSIL-----GMEPMRNATKGCN 229

QY 245 ESQDDVATCSCDCAASC-----PAIARQALDSTF-----YLGQMPGSIV 286
DB 230 ESDVETGTGSCDSCDSIVCGPKQPPPPMPWRINGLDAMVIMVYVAFILVFVFGALL 289

QY 287 LIILCSFPAVVTILLVGRVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQFFQGMGT 346
DB 290 AVWCHRRRYFVSEYTPIDSNIAFSVNSS---DKGEASCCDPLGAADFDDCLRRMFTKGA 345

QY 347 WVASWPLTILVLVSVIPVVALAGLVTELTDPVELWSAPNSQARSEKAFPHQFGPPFR 406
DB 346 FCVRNPCTCIFFSLAFITVCSGLVQVQVTTNPVELWSAPHQARLEKEYFDKFGPPFR 405

QY 407 TNOVILTAPNRSRYRDSLLGPK-NFSGILDLDLLELLEQLERLHLQVNSPEAQRN- 464
DB 406 TEGLIIQANTSVHIIYEPTPADADVFGPPLNKEILHQLVNLQ-----TAISITASYNN 460

QY 465 --ISLQDICVAPLNPNSTSLYDCINSLQYFQNNRTLLLTANQTLMGQTSQV----- 516
DB 461 ETVTLDQICVAPLSPYK---NCTIMSVNLVFNQNSHAVL-----DSQVGDGDFYI 506

QY 517 --DWKDHFLYCANAPLTFKDGATALASCHADYCAPVFPFLAIGGYKGDYSENAALIMTF 574
DB 507 YADYTHFLYCVRAPASLNDTSLHGCPLGTGCGPVPFVPLVGLGQDDQYNNATALVITF 566

QY 575 SLANNYPAGDRLAQAKLWEAFLEMRKAFORMMAGMFQVTFTAERSLDEINRTTAEDLP 634
DB 567 PVNNYNDTERLQANAWKEKFIISFKVNYKN---PNLTISFTAESIEDELNRESNDVF 623

QY 635 IPATSYIVIFLYISLALGSYSSWRVWVOSKATLGLGGVAVILGAVMAAGPFSYIGIRS 694
DB 635 IPATSYIVIFLYISLALGSYSSWRVWVOSKATLGLGGVAVILGAVMAAGPFSYIGIRS 694

DB 624 TVIISYVMEFLYISLALGHIQSCSRLLVDSKISLGIAGIILVILSSVACSLGIFSYMGMLP 683

QY 695 SLVILQVVPVLVLSVGADNIFIVLEQYORLPRRGEPREVIHGRALGRVAPSMCLLSLSE 754
DB 684 TLIVIEVIFPLVAVGVDNIFILVQTYQRDERLQEBTLDQOGLGILGEVAPTMELSFSE 743

QY 755 AICFFLGALTPMPAVRTRFALTSLGLAVILDFLLQMSAFVALLSLDSKRQESRLRDLVCCVK 814
DB 744 TSAFFFGALSMSPAVHTFSLPAGMAVILDFLLQITCFVSLGLLDIKQEKHNLIDLCVR 803

QY 815 PQELPPPGQ-----EGLLLGFFQKAYAPFLHMITRGVLLFLFALFGVLSYMSCHISVG 870
DB 804 GAD---DQCGSHASESYLFFFKYFAPLLLLKDWLRPIVVAVFVGLVSFSAVVNVKVDIG 860

QY 871 LDQELALPKDSYLLDYFLNRYPEVGPVVFVTTLGYNFSSSEAGMAICSSAGCNPFSF 930
DB 861 LDQSLMPNDSYVIANFKSLAQYLSHGGPPVTFVLEEGTNSRRKGQNMVCGMGCDNDSL 920

QY 931 TQKIQVATEPPEQSYLAIPASSWVDDIDMLTP--SSCCRLYISGPNKDKFCPSTVNSLNC 989
DB 921 VQQIFNAELDTYTRVGFAPSSWIDDYFDWVSPQSSCCRLY---NVTHQFCNASVMDPTC 977

QY 990 LKNCMSIT-MGSVRPSVEQFHKYLFWFLNDRPNIKCPKGLAAAYSTSNLTSQGVLA-AS 1047
DB 978 VR-CRPLTPEGKQRPQCKEPMKFLPMLFSLDNPNPKCGKGHAAYGSAVINVGDDTYIGAT 1036

QY 1048 RFMAYHKPLKNSQDYTEALRAARELAANITADLRKPGTDPAFEVFPFYTIITNVFYEQYLT 1107
DB 1037 YFTYTHILKTSADYTDAMKARLIASNIYETWES-KGSD--YRVFPYSVVFVEQYLT 1093

QY 1108 ILPEGLFMLSCLVPTTFAVSCLLGLDLRSLGLNLLSIVMLVDTVGPMLAWDYSYNAVS 1167
DB 1094 IIDDITFNLSVLSIGSIFLTVLVGLCELWSAVIMCITAMILVNMFGVMWLGISLNAVS 1153

QY 1168 LINLVSAGMSVEFVSHITRSFALSTKPTWLERAKEATISGSAVFAGVAMTNLPGIILVL 1227
DB 1154 LVNLVMSCGISVEFCSHITRAFTMTSGRSVRSAREALAHMGSSVFSGITILTKFGGIVVL 1213

QY 1228 GLAAQALQIIFFFRLNLLITLLGLLGLVFLPVLISYVGPDPVNP 1272
DB 1214 APAKSIQEIFEYFRMYLANVLLGATHGLIFLPLVLLSYLGPVSNKA 1258

RESULT 2

T05663

hypothetical protein F22113.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T05663

R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15420

A:Accession: T05663

A:Molecule type: DNA

A:Residues: 1-1055 <BEV>

A:Cross-references: UNIPROT:Q9SVF0; UNIPARC:UPI00000A52EC; EMBL:AL035539

A:Experimental source: cultivar Columbia; BAC clone F22113

C:Genetics:

A:Map position: 4

A:Introns: 24/3; 60/3; 99/3; 150/1; 193/1; 216/3; 278/3; 297/2; 336/3; 364/3; 396/3; 411/3

026/3

A:Note: F22113.120

Query Match 20.2%; Score 1394; DB 2; Length 1055;

Best Local Similarity 28.5%; Pred. No. 2.6e-89;

Matches 377; Conservative 221; Mismatches 385; Indels 340; Gaps 41;

QY 36 YDECGNPGLSGSLMTLSNVCLSNTPARKITGDHLI--LLQKICPRLTYGTNTQACCSA 93

DB 2 YDICHRSD--GKVL-----NCPYASPS--IQDELFSAKIQSLCPTI-----SGNVCCTE 48

QY 94 KQLVSLASLSITKALLTRCPACSNFVNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAV 153


```
Db      515  ---TQANWEER-LEEYLLDUKPEGL-RISFNTIEISLEKLN--NDISTVAISYLM 567
Qy      643  IFLYISLALGSYSSRVNVDKATIGLGGVAVLVGAVMAAMGFFSYLGIRSLVTLQV 702
Db      568  MFLYATWALRRKDGKTRLL-----LGISGLLIVLASIVCAAGFLTLFLGKSTLIIAEVI 621
Qy      703  PFLVLVSGADNIFIVLEYQV-LPRRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLG 761
Db      622  PELILAIGIDNIFLITHEIDRNCQPEYSIDQKIISAIGRMSPSILMSLLCQTGCFLIA 681
Qy      762  ALTPMPAVTEALTSLGVLDFLOMSAFVALLSLDSKQESRLDVCCKVPQELPPP 821
Db      682  AFVTPAVNFAIYSTVIFNGVQLTAYVLSLYEKRSNYKQIT----- 728
Qy      822  GQEGILLGFFKAVAPFLLHMTIRGVILLFLALFGVLSYSMCHISVGLDQELALPKDS 881
Db      729  -GNEETKESFLKTFYFKMLTQ--KRLIIIIISAFMFFTSVLVFLPELOFGLDQTLAVPQDS 784
Qy      882  YLLDPLFLNRYFEVGAPVYVTTTLGYNFSSBAGMAICSS-AGCNPFSTQKIQVATRF 940
Db      785  YLVDFPKDYSLNNGPPVYVYVVK-NLDTKRQNOQKICGKFTTCERDSLAVLE---QE 840
Qy      941  PQSYLAIPASSWDDPDLWLPSS---CCRLYISGNKDKFCPSVNSLNCNCKMSITM 998
Db      841  RHRSTITEPLANWLDYFMLPQNDQCRL---KKGTDVPCPPSPSRRC-ETCFQ--Q 894
Qy      999  GSVR-----PSVEQFHKVLPFLNDRPNKCPKGLAAYSTSNLSDGVVLASRFMAY 1052
Db      895  GSWNYNMSGFFGKDFMEVLSWIN-APSDPCLGRAPYSTAL-VYNETSVASVFRTA 952
Qy      1053  HKPLKNSQDYTEALRAARELANIATADLRKVPGTDPAPFVPPYTTNVFYQYLTLPEG 1112
Db      953  HHPLRSQKDFIQAY-----SDGVRISSSPPELDMFAYSPPYIPFYQYQTLGPLT 1001
Qy      1113  LFWLSICLVPTFAVSCLLGLDRLSLNLLSIVMLVDTVGFMAWDISYNAVSLINLV 1172
Db      1002  LKLGSAIILIFPISVFL-QIRSSFLALVVTMIIVDIGALMALLGISLNAVSLVNI 1060
Qy      1173  SAVGMSVEFVSHITSFAL---STKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGL 1229
Db      1061  ICVGLGVEFCHVHSFTVVPSTKTDANSRVLSYNTIGESVIGKILTKFYGVCVLP 1120
Qy      1230  AKAQILQIPFFRLNLITLLGLLHGLVFLPVLISYVG 1266
Db      1121  AQSKIPDFYFRMFWFTLIIVAALHALLPLPALLSLFG 1157

RESULT 5
T30172
transmembrane protein patched - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30172
R:Goodrich, L.V.; Johnson, R.L.; Milenkovic, L.; McMahon, J.A.; Scott, M.P.
Genes Dev. 10, 301-312, 1996
A:Title: Conservation of the hedgehog/patched signaling pathway from flies to mice: indu
A:Reference number: Z20752; MUID:96176226; PMID:8595881
A:Accession: T30172
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1434 <GO>
A:Cross-references: UNIPROT:Q61115; UNIPARC:UPI0000021C02; EMBL:U46155; NID:g1181884; PI
A:Experimental source: clone M2, M9
C:Superfamily: Drosophila membrane protein patched

Query Match      9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 9e-37;
Matches 283; Conserved 171; Mismatches 420; Indels 363; Gaps 44;

Qy      356  LVLSVLPVVALAAGLVFTLTDDPVELWSAPNSQARSEKAFHQHGFPPFRN-QVILTA 414
Db      88  LVVGLLIFGAFVGLKAAANLETNVELVVEGGRVSRLENTYTRQKIGBEAMFNPLMIQT 147
```

```
Qy      415  PNRSYRYSLLGLPKNFSGIILDLLELE--LQERLRLHQLQVMSPEAQRNISLQDICY 472
Db      148  PKEEG-----ANVLTRALLQHLDSALQASRVHVVMYN---RQMKLEHLCY 190
Qy      473  AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLLLTANQTLMGQ----- 512
Db      191  KSGELITETGYMDQIIEYLYPCLIIITPLDCPWEKALQSGTA--YLLGKPLRLWTNFDPL 248
Qy      513  -----TSQVDWKDHFY-----CAN-----APL----- 530
Db      249  EFLBELKKNYQVDSWEEMLNKAEGVGHGMDRPNCLNPDPCPATAPNKNSTKPLDVALV 308
Qy      531  -----TFKQGT-----ALALSCMADYGAPVFPFLAIGGYKKG 562
Db      309  LNGGCGGLSRKYMHWQEBELIVGGTVKNAITGLVSAHALQTMFLQMTPKQMYEHFQY--- 365
Qy      563  DYSEBALIMTFSILNYPAGDPLRLAQAKLWEBAFLBEMRAFORRMAGMFQVTFTRBSLE 622
Db      366  DY-----VSHINWNE-----DRAAAIILSAWQRTYVEVHVQSVAPNSTQKVLFPFTT--TTLD 414
Qy      623  DEINRTTAEDLPIFATSYIVIFLYISLALGSYSSRVNVDKATIGLGGVAVLVGAVNA 682
Db      415  DILKSPFSDVSVIRVASGYLLMLAYACLTM---LRWD---CSKQGAIVGLAGVLVALSVAA 469
Qy      683  AMGFPSYLGIRSSVLQVVPFLVSVGADNIFIVLEYQRLPRRPGPREVHIGRALGR 742
Db      470  GLGLCSLIGISFNAATTQVLPFLALGVGDVDFLLAHAFSETGQNKRIIPEDRTCECLAR 529
Qy      743  VAPSMMLCSLSEAICFFGALTPMAVTRFALTSLGLVLDLFLQMSAFVALLSLDSKQ 802
Db      530  TGASVALTSISNVTAFPMALTIPIPALRAFSLQAAVVVFNFAVLLIIPALLSMDLYR 589
Qy      803  EASRLDVCCK-----VKQEL-----PPCQGGEL----- 827
Db      590  EDRLLDIFCCFTSPCVSRVIVQEPQAYTEPHSNTRYSPPPPYTSHFAETHITWQSTVQ 649
Qy      828  ----- 827
Db      650  LRTEYDPHTHYVYTTAEPRESEISVQPVTVTQDNLSQCSPESTSTDRLLSQSDSLHCL 709
Qy      828  -----LLGPFQAYAPPFLHMTIRGVVLLFLALFGVLSYSMCHISVGLDQELALPK 879
Db      710  EPPCTKWTLSFPAEKHYAPFLKPKRAKVVVILLFLGLLGVSLYGTTRVRDGLDLDIVR 769
Qy      880  DSYLLDYLFLNRYFEVGAPVYVTTGLGYNFSSBAGMAICSSAGCNPFSTQKIQYAT- 938
Db      770  ETREYDFIAQFKYFSP-----YNM-----YIVTKADYPNI 801
Qy      939  -----EPPEQSYLAIPASS-----WVDDFIDWL-----TPSS--- 965
Db      802  QHLLYDLHLKHSFNVKYVMLEENKQLPQMWLHYFRDLWLOQLQDAPDSDMETGRIMPNYKN 861
Qy      966  -----CCRLYISGNKDKFCPSVNSLNCNCKMSITMGSVRSVRSQFHKYL--PWFLN 1017
Db      862  GSDGVLAYKLVAVQGRDK--PIDISQIT--KQRLVDADGDIINPSA--FYIYLTAWSN 915
Qy      1018  D-----RPNKCPKGGAAAY--STSVNLTSQGVLASRMAVYHKKPKNSOD 1061
Db      916  DPVAAASQANIRPHRPBWHDK---ADYMPETRURIPAAEPDIEVAQPFYFYLNGRLDTS 972
Qy      1062  YTEALRAARELAANITA-DLRKVPGTDPAPFVFPYTTNVFYQYLTLPEGLFPLMSLCL 1120
Db      973  FVEALEKRVICNNYTSGLSSYPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
Qy      1121  VPTFAVSCLLGLDRLSGLLNLLSIVMLVDTVGFMAWDISYNAVSLINLVASVGMSTV 1180
Db      1023  ACTFLVCAVFLINPWTAGII--VMVLAMTVELFGMGLIGIKLSAVPVVILLIASVIGVE 1081
Qy      1181  FVSHITRSP--AISTKPTWLERAKEATISMGSAVPAGVAMTNLPGLVGLAKAQLIQIF 1238
Db      1082  FTVHVALAFLTAIGDN---HRAMLALBHMFAVLDDG--AVSTLLGLVLMAGSEFPIVRY 1137
```

QY 1239 FFRNLITLGLLGLVFLVILSYVG--PDVNPALALEQ---KRAEEAVALVAVASCP 1293
DB 1138 FFAVALITLVLGLVLLVLLVLLSFFGCPCEVSPANGLNRLPTSPSPSPVVRFAVPP 1197
QY 1294 NHPKRSV--TADNIYNHVSFEYSI-----KGAG 1319
DB 1198 GHTNNGSDSSSEYSSQTTVSGISEELRQYEAQAG 1234
RESULT 6
T18538
patched protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18538
R:Marigo, V.; Scott, M.P.; Johnson, R.L.; Goodrich, L.V.; Tabin, C.
Development 122, 1225-1233, 1996
A:Title: Conservation in hedgehog signaling: induction of a chicken patched homolog by S
A:Reference number: Z18958; PMID:96205046; PMID:8620849
A:Accession: T18538
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1442 <MAR>
A:Cross-references: UNIPROT:Q90693; UNIPARC:UPI00001328B4; EMBL:U40074; NID:G1335850; P1
C:Gene: PTC
C:Superfamily: Drosophila membrane protein patched
C:Keywords: transmembrane protein
Query Match 9.1%; Score 628.5; DB 2; Length 1442;
Best Local Similarity 23.6%; Pred. No. 1.6e-35;
Matches 284; Conservative 165; Mismatches 421; Indels 335; Gaps 47;
QY 365 ALAAGLVFTELTDPVELWSPNSQARSEKAFDQHFQFFFTN-QVILTAPNRSRYD 423
DB 111 AFAVGLRAANLETNVEELWVEGVRVRELNYTRQKIGEEAMFNQPMIQPOEDG---- 166
QY 424 SLLGPKNFGILDLLLELLE--LQERLRLQWSPQAQRNLSLQDICYAP----- 474
DB 167 -----TNVLTTEALRQHLDSALQASRVHYMYN-----RQWKLEHLCKYKSGBLITEA 213
QY 475 --LNPDNTSLYDCINSLLQYFQNNRTLLLTANQTLMGQ----- 512
DB 214 GYMDQIIIEYLCIITPLDCFWEGAKLQSGTA--YLLGKPLQWINFDFLEELKKI 271
QY 513 TSQVDMKDFLY-----CAN-----APLTF--KQGT-----ALALS--CMADY 546
DB 272 NYQVESWEEMLNKAIEVGHGYMDRCPCLNPADPCPITAPNKNSTKPLDVALVLSGGC---Y 328
QY 547 GAPVFPF-----LAIGG-----YKGDYSEAEALIMTFSL-----NNYP----- 580
DB 329 GLSRKYMHWQELIIGTVKNSGKLVS--AQALQTMFQMLTPKQMYEHFGYEVSHINW 387
QY 581 AGPRLAQAKLMEAELEEMRAFORRMAGMFQVTFTEARSLDEINRTTAEDLPFATSY 640
DB 388 NEDKAAALEAMQRMVEVHVQSVQNSQKVLSTFT--ITLDDILKSPSDSVIRVASQY 446
QY 641 IVIFLYISIALGYSYSSWSRVMDSKATLGLGVAVVLGAVMAAMGFYSLGRSSVLILQ 700
DB 447 LMLAYACLTM-----LRWD--CAKSOQAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQ 501
QY 701 VVPLVLSVGCADNIFIVLEYQRLPRRPGEPREHVHIGRALGRVAPSGMLCSLSEACFFEL 760
DB 502 VLFPFLALGVGDVDFLLAHAFSTGQNKRIPFEDRTGECUKRTGASVALTSISNVTAFPM 561
QY 761 GALTMPAVRTFALTSGLAVIDFLQMGSAFALLSLDSKQBSRLDVCCC----- 812
DB 562 AALIPALRAFSLQAQVVVVFAMVLLIFPAILSNLDVRRDRDLDFCCFTSPCVTR 621
QY 813 ---VKPO-----ELPPQGGEGE----- 827
DB 622 VIQIEPOAYANDNICYSSPPPYSSSHFAHETQITWQSTVQLRTEYDPHTQAYYTTAEPR 681

QY 828 -----LLGFFQKAYAP 838
DB 682 SEISQVPTVTQDLSQCPESASSTRDLLSQSDSSVHCLPEPCTKWTLSSTFAEKHYAP 741
QY 839 FLHWHITRGVLLFLALFQVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 898
DB 742 FLKPKAKVVIIFLGLGLLSLYGTTRVRDGLDLTDIVPRDREYDFIAAQPKYFSF-- 799
QY 899 PTVFTVLGYNFSSEAGMNAICSSAGCNFPSTQKIQYAT-----EPPEQSYLAI 948
DB 800 -----YNN-----YIVTKADYPNVQHLHLSFSNVTVLL 833
QY 949 PA-----SSWDDDFIDWL-----TFSS-----CCRLYISGPNKD 977
DB 834 EGRQLPKWLHLFRDWLQGLQDAFDSQWETGKITYSYKNGSGDDAVLAYKLLVQTGNRA 893
QY 978 KFCPSTVNSLNCNKNCSITMGSVRSVQFQHKYL-PWFLND-----RPN 1022
DB 894 K--PIDISQLT--KQRLVDADGIIINPNA--FYIYLTAWVSNDFVAYAAQANIRPHPEW 947
QY 1023 KCPKGGIAAY--STSVNLTSDGOVLASRRFAYHKKLNKSNQDYTEALRAARELAANITA-D 1079
DB 948 VHDK---ADWPETRLRIPAEFIEYAOFFFYLNGLRETSDFVEAIEKVRACNNYISLG 1004
QY 1080 LRKVPGTDPAPFVPPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGL 1139
DB 1005 IASYPNG-----YPP-----LFWEQYIGLRHMLLSISVVLACTFLVCALFLNPNWTAGI 1054
QY 1140 LNLISVIMLVDTVGFMALWDISYNAVSLINLSAVGMSVEFVSHITRSP--AISTKPTW 1197
DB 1055 I-VVVALMTVELFGMGLIGIKLSAVPVVILIASVGIVGVEFTVHIALAFULTAGDKN-- 1111
QY 1198 LERAKETISGSAVFAVAVMTNLPGLVLGLAKAQIQIPIFFRLNLIITLLGLLHGLV 1257
DB 1112 -RRVLALEHMFVLDG-AVSTLLGLVMLAGSEDFIVRYFAVLAITLILGLVGLV 1169
QY 1258 LPVILSYVG--PDVNPALA---LEQKRAEAAVAAVMVASCNHPSRVSTADNIYNHSE 1312
DB 1170 LPVLLSFFGYPPEVSPACGRNRLPTSPPEPPPSIVREALPGHTNNGS--DSSDSEYSSQ 1227
QY 1313 GSIKG 1317
DB 1228 TVVSG 1232
RESULT 7
S44797
F09G8.4 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S44797
R:Anderson, K.
submitted to the EMBL Data Library, February 1993
A:Description: Sequence of the C. elegans cosmid F09G8.
A:Reference number: S44796
A:Accession: S44797
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-915 <AND>
A:Cross-references: UNIPROT:P34389; UNIPARC:UPI000013B859; EMBL:L11247; NID:G156280; P1D
C:Gene: C
A:Introns: 51/1; 214/3; 255/3; 326/3; 382/2; 409/3; 455/3; 505/3; 569/2; 702/3; 836/3
Query Match 8.7%; Score 600; DB 2; Length 915;
Best Local Similarity 22.7%; Pred. No. 8.9e-34;
Matches 236; Conservative 158; Mismatches 422; Indels 224; Gaps 37;
QY 10 LWMALLRLAQSEPTTHIQPGYCAFYDECGKNPBLSGSLMTLSNVCLSNTPARKITGD 69
DB 14 VLFLLLIHLALCQ-----AKVM-TECDGEDSNHPPCKTNKSTYLPITVTRSLNPT 64
QY 70 HLILLOKICPLRYTGN--TQACCSAKOLVLEASISITKALLTRCPACSDNFVNLHCHNT 128

Db 643 -----DDNGTMVKSFRFILGMDKLVTTMDQTDATMSFREVAARM---- 681

Qy 1081 RKVPGTDPAFVFPYTTNVFQYLTITLPGELFMLSCLVPTTAVNSCLLLGLDLRSGLL 1140

Db 682 -----PEFNVTTFMPIMFTDQYIIIPNTQNIILALVMIVIAVLFPIQPMCS-LW 733

Qy 1141 NLLSIVMLTVDVGMALWDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200

Db 734 VALACASIDFGVIGTWTUWGNLDAISMITIIMSIGFSDYSAHIAVGVVSRDTAAGR 793

Qy 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLLITLLGLLHGLFLPV 1260

Db 794 VKEALSALGWPLSQG-ANSTIIAVSVLADIPAYMI-VTFEKTVVLISLGLLHGLFLPV 851

Qy 1261 ILSYV-----GPDVNPALALEOKRAEA-----VAAMVAVSCPNHPSRVS 1300

Db 852 LLSIFVRGCCIIIPSPHPGHPSAQKIEQIRIAAISSSPLDLRTVAPLRASSPISFPHPHLE 911

Qy 1301 -TADNIVYNHSPGSIK 1316

Db 912 YDESPTVHNRKNSIK 928

RESULT 9

T13952

membrane protein ptch2 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T13952

R:Motoyama, J.; Takabatake, T.; Takeshima, K.; Hui, C.

Nature Genet. 18, 104-106, 1998

A:Title: Ptch2, a second mouse patched gene is co-expressed with Sonic hedgehog.

A:Reference number: Z17830; MUID:98122566; PMID:9462734

A:Accession: T13952

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1182 <MOT>

A:Cross-references: UNIPROT:Q35595; UNIPARC:UPI0000001587; EMBL:AB010833; PIDN:BAA24691.

A:Experimental source: strain BALB/cCrSlc

C:Gene: ptch2

C:Superfamily: Drosophila membrane protein patched

C:Keywords: transmembrane protein

Query Match 8.4%; Score 583.5; DB 2; Length 1182;

Best Local Similarity 24.2%; Pred. No. 1.8e-32;

Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

Qy 330 SFSTHTLLGQ-----FFQG-----WGTWVASWPLTILVLSVIPVVALAAGLVFTELT 376

Db 20 SSAPHILAGSLQAPLMIRAYFQGLLSLCRIQKCHGVFLGLVAFGALALGLRVAVIE 79

Qy 377 TDPVELWSPNSQARSEKAPHDQHG-PFRTNQVILTAPNRSSVRYSDLSLLGPNFSGI 435

Db 80 TDLEQLWVEGSRVSQELHYTKELGEEAAYTSQMLI-----QTAHQEGGNVLTPE----A 131

Qy 436 LDLDLLELLELQERLRLHQLVMSPEAQRNLSQDICY---APLNPDN-----TSLYDC 486

Db 132 LD-----LHLQAALTASKVQVSLYKGSWDLNKICYKGVPLTIENGMIERIEKLPFCV 184

Qy 487 INSLLQYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT----FKQGTALA 539

Db 185 ILTPLDCFWGAK---LOGGSAYLPGRPDIQWNTNLPQQLLELGPFPASLEGFRELDKA 241

Qy 540 LSCMADYGAPVF-----PFLA---IGGYKGDYVS----- 565

Db 242 QVGQAVYGRFCLDPDDPHCPFPAPNRHSRQAPNVAQELSGGCHGFSHKFMHWQOEILLGG 301

Qy 566 -----EAEALIMTF-----SLNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAG 609

Db 302 TARDLOQLRAEALQSTFLLMSPRQYEHFRGDYQTHDIGNSEEQASWVLQAWRRFVQ 361

Qy 610 MFQVTFTAERSLE-----DEINRTTAEDLPFATSYIV--IFLYTISLALGSYSWS 658

Db 362 LAQEALPANASQOIHAFSSTTLDDILRAFSE-----VSTTRVVGGLMLLAYACVTMLRWD 417

Qy 659 RVNVDSKATIGLGGVAVVLGAVMAAMGFPSYILGRSLVILQVVPVVLVLSVGDANFIFV 718

Db 418 --CAQSQGAVGLAGVLVALAVASGLGCLGALLITNAATTQVLPALGIGVDDIFLLA 475

Qy 719 LEYQRLPRRPEGEPREVHIGRALGRVAPSMILCSLSBAICFFLGALTPMPAVRTFALTSG 778

Db 476 HATKAP--PDTPLPERMGBCLRSTGTSVALTSVNNMVAFMAALVPALRAFSLQAAI 533

Qy 779 AVILDFLQMSAFVALLSLDSKQEAASRLDVCC-----VKPQE----- 817

Db 534 VVGCNFAAVMLVPAILSLDRHRQRDLVLCFCFSPCSAQVIQMLPQELGRVAVPGI 593

Qy 818 -----LPP-----PG-----QSEG-- 826

Db 594 AHLTAVTQVATHECASSQHVVTILPQAHLLSPASDPLGSELSPGSGSTRDLLSQEBSGT 653

Qy 827 -----LLLGFFQKAYAPFLHWTIRGVVLLFLALFGVSLYSNMCHISVGLDQ 873

Db 654 PQAACRELLCAHWTLAHFARYQFAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713

Qy 874 ELALPKDSYLLDYFLNRYFEVGAIVYFTTILGYNFS-SEAGMNAICSSAGCNPSTQ 932

Db 714 TDVVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDYAHSQRAL-----FDLHQ 761

Qy 933 KIQVATEFPQSILAIIPAS-----SWDDF-----IDMLTPSSCCRLYISGPNK 976

Db 762 RFSSL-----KAVLPPPTAQAPRTWLHYRSWLOGIOAAFPDQWASGRITCHSYRNGSD 816

Qy 977 D---KFCPSVTNSLNCIKMCSITM-----GSVRPSVQFHKYLPWFLNDRNRIKCPK 1026

Db 817 GALAYKLLIQTGNAQEPDLFSQLTTRKLVKDEGLIPP--ELFYMGLTVVWSSDPL----- 869

Qy 1027 GGLAA-----YST---SVNLTSQGVLASREPMAYHKPLKNSQDYTRALR 1067

Db 870 -GLAASQANFYPPPPPEWLHDKYDTTGNLRIPAAQPLEFAQFPFLHGLQKTADFVEAIE 928

Qy 1068 AARELAANI-TADLRKVPGTDPAPFVFTYITNVFYEQYLTILPEGLFMLSIC--LVPTF 1124

Db 929 GARAACTEAGQAGVHAYPSGSPF-----LFWEQYLG--RRCPFLAVCILLVCTF 976

Qy 1125 AVSCILLGLDLRSGLNLLSIVMLVDTVGPMAWLDISYNAVSLINLVSAVGMSEVFSVSH 1184

Db 977 LVCALLLSPTWAGLI-VLVLAMTVELFGIMGLGINKLSAIPVIVLASIGIGVEFTVH 1035

Qy 1185 ITRSPAISTKTWLERAKEATISMGSAVFAGV---AMTNLPGILVLGLAKAQLIQIFPFR 1241

Db 1036 VALGELTSHGSRNLRAA-----SALEQTFAPVTDGAVSTLLGLMLAGSNFDFIIRYFV 1090

Qy 1242 LNLITLTLGLLHGLVFLPVILSVGP 1267

Db 1091 VLTVTLTLGLLHGLLLEPVLISILGP 1116

RESULT 10

T36746

hypothetical protein Y39A1B.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T26746

R:Wall, M.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20258

A:Accession: T26746

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1003 <WLL>

A:Cross-references: UNIPROT:Q9XXR9; UNIPARC:UPI00000782FA; EMBL:AL021482; PIDN:CAAL6339

A:Experimental source: clone Y39A1B

C:Gene: CESP:Y39A1B.2

```

A;Map position: 3
A;introns: 89/2; 126/3; 167/2; 194/2; 232/1; 266/3; 295/2; 323/2; 363/1; 519/3; 696/3; 7
Query Match      8.3%; Score 573.5; DB 2; Length 1003;
Best Local Similarity 22.8%; Pred No. 7.3e-32;
Matches 236; Conservative 198; Mismatches 376; Indels 225; Gaps 42;

Qy 339 QFGGWTG---VASWPLTLTILSVIPVVALAAGLVTE--LTTDPVELWSAPNSQARSE 393
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 9 RFAHAGAYSGVVVHPHPPFFIIPILITAGLSTGLLRHEQAFMKDELELYTPTDQARKE 68
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 394 KA-----PHQHGFPPFRNQ-----VILTAPNRSYRYSLLLPKPNFSGILDLDL 440
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 69 LSQDLHLPHINDSPFYATRYDIRRAGYIIVT---NOEEDGDILNPLVHMSANQLWS 123
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 441 LLELLEQLERLHLQVMSPEAQRNISLQDIC-----YAPLNPNTSLDYDCCI 487
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 124 IVQSLTVED-----EDDRINYPSCVFPPIPPPEFSKALHSLFAPNMTTPEICV 173
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 488 -NSLQYFQNNRTLLLTANQTLMGOT-----SQVMDKHFLYCANAPLTPKPGTALAL 540
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 174 SNPLVEIFK-----LLLVSDRSFLNRSIDBWLTSQI-----SDAIQPSDGGMTHL 218
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 541 SCHWADICAPVPPFLAIGY---KGKDYSEBALIMTFSLNNYPAGDPRLAQAKLWEAEFL 597
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 219 -----LGGVTLDDDKRTAGAKAMLLPYALRH---SSDDEDVWAKEWVRILA 261
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 598 EEMRAFQRRMAGFQVTFTAERSLDEBINTTAAEDLPFATPSYIVIPFLYISLALGSYSW 657
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 262 DFLLOYD---SPIIRASWWTYETLAEASDRQLQIHMLLPCFVCVSIPTACCCVFSW 317
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 658 SRVMDSKATGLGGVAVLVGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIF 717
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 318 RR-----SRPWLAIIGGVISAAMATASAVIGILLAGYGMTSVAYS-MPFIVSFVGVNDVFIL 372
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 718 VLEQRLPRPGEPREVHIGALGRVAPSMLLCSLSEACFFLGLATPMPAVRTFALTSG 777
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 373 LSAWRSTST--ETLEHRMKETFADAAVSITVTLTDLISFGVGCATPFPVSQMFCAVAV 430
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 778 LAVILDPQLQMSAFVALLSDSKQBASRLDVCVV-----KQELPPPCQGBG----- 826
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 431 AAIVFTYIYQITFFAAVWVNTNRREINN---HCIFHKLKQDTLPEKIAAQDSRPFK 486
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 827 --LILGPFQKAYAPFLHMTITRGVLLIFLALFGVSLYSMCHISVGLDQELALPKDSYLL 884
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 487 NTILAQFRITYSDFLNLPLVRILVITFCVILGVASGCTKVYKLGLENDLLPENSYK 546
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 885 DYFLFLNRYF-EVGAPVYFVTTIGYNPSS-EAGNNAICSSAGCNPNFTQKIQYATEPPE 942
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 547 RTLMAAEKYPFSDYSGSLH---VMMYNLSEVDVAPRKIWN-----VLEKEVELYE 592
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 943 QSVLAIPASSWVDFFIDW-----LTPSSCCRLYISGPNKDKFCPTVNSLNCNLCMS 995
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 593 HTFETASSDNLRTFLAFVQAGLLITPE-----NFVILKX----- 629
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 996 ITMGSRVPSVEQHKYLPFWFLNDRPNIKCPKGLAAYSTSVNLTSDQVL-ASRFMAYHK 1054
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 630 -----VFLSQPQF-----AKYNRDVVLTEDEHLEASRI----- 658
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1055 PLK-----NSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITVTFVEQYUJLIP 1110
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 659 PVQLRHVGSQANQRAMLFRRLAE--TSEIQ-----TGVYADFPQ-----FAEQYNAVLP 706
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1111 EGLFPMLSLCLVPPIFAYSCIALGLDRLSGLLNLSIVMLVDTVGFMAWDISVNAVSLIN 1170
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 707 GTLSSTIATAGVAVVAVSLIILIPFVAS-LWVSFISVINSIGILGPTFWFSVRLDIFSMVT 765
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1171 LVSAVGMSVEFVSHITRSFAISTKPTWLERAKETATISMGSAVPAGVAMTNLPICILLVGLA 1230
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 766 IVMSIGFCDVFAHLVNFYAKQWMDGSEMRNALYAVGAPILMS-ATSTIGVSPMAASA 824
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1231 KAQLIQIFPRLNLLITLLGLLHGLVFLPVILS--YVGPDVNPALAEQKRAEEAAVAV- 1287
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 825 ESYVFR-SPLKTMILVILGALGLVILPVLSSMFYCGG-----SSKAKEHIDAVD 875
 Qy 1288 -MVASCPNHPRSVST 1301
 Db 876 QKQAQYNNPARYAS 890
 RESULT 11
 T18291
 patched protein - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T18291
 R:Concorder, J.P.; Lewis, K.E.; Moore, J.W.; Goodrich, L.V.; Johnson, R.L.; Scott, M.P.
 Development 122, 2835-2846, 1996
 A:Title: Spatial regulation of a zebrafish patched homologue reflects the roles of soni
 A:Reference number: 218860; MUID:96379744; PMID:8787757
 A:Accession: T18291
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1220 <CON>
 A:Cross-references: UNIPROT:Q98864; UNIPARC:UPI0000132883; EMBL:X98883; PIDN:CAA67386.1
 C:Genetics:
 A:Gene: ptc1
 C:Superfamily: Drosophila membrane protein patched
 Query Match 8.3% Score 572; DB 2; Length 1220;
 Best Local Similarity 22.4%; Pred. No. 1.2e-31;
 Matches 270; Conservative 174; Mismatches 431; Indels 328; Gaps 43;
 Qy 321 KGTSLSDKLSFSTHTLLGQPFQGGTGWASWPLTILVLSVIPVVALAAGLVFTLTDPV 380
 Db 51 KGRVAGQKAPLWTRARFOALFSLGCHIQRHCGKVLPIGLLVFGALS VGLRAVALETDI 110
 Qy 381 ELMSAPNSQARSEKAFDQHFQ-PFFTNQVILTPNRSSYRYDSLLGPKNPSGILDLD 439
 Db 111 KLVWAGSRVSKELRYTKRQGESEVTSQMLIQPKQEG-----TNLTQE 157
 Qy 440 LLELELEQLRLRHLOWSPSAORNISLQDICYAPLNP--DNT-----SLYDCCINSL 490
 Db 158 AL--LLHLEALSGKQVSLYKSGWDLNKLCKFGVPIIENVMIERMIDKLPQMWTP 215
 Qy 491 LQYFQNNRTL-----LLLTANQ--TLMGQTSQVDWKDHFLYCA 526
 Db 216 LDCFWEGSKLGGSAYLPGMPDIOQNNLDPLKMEELSQTSLSEGFREMLDKAQVGHAYM 275
 Qy 527 NAPIPTFKDGTALALSCMADYCAP-----VPPFLAI--GGYKG-----561
 Db 276 NRPLCDPSDT----DC--PHSAPNKPQWPNIAAEIQGGCHGFGSKPMHWQEBELILGER 329
 Qy 562 -KD-----YSRAEALIMTFSL-----NNYPAGDPRLAQAKL-----WEBAFLKE 599
 Db 330 VKDSQNALQSABALQTFWLMSPKQLYEHFKDDTEIHDINWEDKATILESWORKFVEV 389
 Qy 600 MRAP--QRRMAGCPQVFTTAERSLEDEINRTTAEIDLPIFATSYIVFLYISLALGSYSW 657
 Db 390 VHGSIPQNSSNVYAFSTT--TLNDIMKSFSDVSIVRAGGYLLMLAYACVTM--LRW 443
 Qy 658 SRVWDSKATLGGVAVIVGAVMAAGPFSYLGIRSLVILQVVPFLVLSVGADNPIF 717
 Db 444 D--CAKSGQAGVGLAGVLLVALSVAAAGLGLCSLGSFNAAATQVLPSLALGIGVDDMFL- 500
 Qy 718 VLEVQRLPRPGRPREVHIGRALGRVAPSMLLCSLSBAICFFLGALTPMPAVRTFALTSG 777
 Db 501 -LGHSFTETRSNIPFKERTGDCLRTQTSVALTSVNNNIAFFMAALVPIPALRAFSLQA 559
 Qy 778 LAVILDLLQMSAFVALLSDSKRQEASRLDVCCK-----VKPQE-----817
 Db 560 VVVVFPMALLFPALISLDLHRRKDLIDLCFFYSPCSSRVIQIQPQELSANDNHQ 619
 Qy 818 -----LPPQEQ-----823
 Db 620 RAPATPYTGSTITTSHTTTTVOAFTQCDAAAGOHIVTLPPTSQISTPPSSMWLSTPTP 679

QY	824	-----GEGL-----	LLGFFOKAYAPFLLHWTR	846
DB	680	TTDPYGSQVTTSSSTRDLQAQVEPKREGRECVLPFFRWNLSFFAREKAPLKKPETK		739
QY	847	GVVLLFLALFGVSLYSNCHISVGLDQELAPKDSYLLDYLFLNRYFEVGAPVYFVTL		906
DB	740	TVVVVVALLSLSLYGTWHDGLYLDIVPRDTQVEYFTAQPKYFSP-YNNMLVTMD		798
QY	907	GVNPSSEAGMNAICSSAGCNFFSTQKIQYATEPEQSULA-----IPASSWDDFDW		960
DB	799	GFYAR-----SQRLQLHNAFNSVKYVVKDGNHKL-P-RMWLHYFQDW		841
QY	961	LTPSSCC-----RLYISGNPKDKFCPSTVNSLNCNKC		993
DB	842	LKGLQATFDADWEAGKITDYSYRNGTDCALAYKPLIQTGSKGPPFNSQLTSRRLVDG-		900
QY	994	MSITWGSVRPSVEQHKYLP-WFLNDRPNIKCPKGLAA-----YST---		1034
DB	901	-----DGLTPPEV--FYIYLTWVNSD-----PLGYAASQANFYHPHREIHDKYDTGE		948
QY	1035	SVNLTSDGQVLASRFMAVHKPKNSQDYTEALRAARELAAN-ITADLAKVPGTDPAPEVF		1093
DB	949	NLRIPAAEPLEPAQPFYFYLNGLRQASDFEAIESTICEBFMRQGIKNYPNG-----Y		1002
QY	1094	PVTITNVFEQYKTLTLPGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTV		1153
DB	1003	PF-----LFEQYIGLRHWFLLSISVVLACTFLVCAILLNPNWTAGVI-VFILPMWTELV		1057
QY	1154	GFMAWDISYNVSLINLVSAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSA		1211
DB	1058	GIMGLIGIKLSAIPVILLASVIGVEFTVHIALGFLTAIGDRNT---RSAVAMEHMEAP		1114
QY	1212	VFAGVAMTLPGLVLGLAKAQLIQIFPRFLNLLTLGLLGLVFLPVILSYGVP--DV		1269
DB	1115	VIDG-AISTLLGVMLAGSEFDMRYPPFAVLAITLLGILNGLVLLPVLLSLMGPPAEV		1173
QY	1270	NPA 1272		
DB	1174	VPA 1176		
RESULT 12				
T29590				
hypoetical protein F55F8.1 - Caenorhabditis elegans				
C:Species: Caenorhabditis elegans				
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004				
C:Accession: T29590				
R:Gatting, S.; Scheet, P.; Kemp, K.				
submitted to the EMBL Data Library, November 1996				
A:Description: The sequence of C. elegans cosmid F55F8.				
A:Reference number: Z20647				
A:Accession: T29590				
A:Status: preliminary; translated from GB/EMBL/DDBJ				
A:Molecule type: DNA				
A:Residues: 1-889 <GAT>				
A:Cross-references: UNIPROT:P91346; UNIPARC:UPI000017BA20; EMBL:U80447; PIDN:AAB37812.1;				
A:Experimental source: strain Bristol N2; clone F55F8				
C:Genetics:				
A:Gene: CESP:F55F8.1				
A:Map position: 1				
A:Introns: 36/3; 66/3; 98/3; 149/2; 200/3; 240/2; 279/1; 358/1; 481/2; 569/3; 606/3; 697/3				
Query Match				
Best Local Similarity 7.8%; Score 539.5; DB 2; Length 889;				
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;				
QY	401	FGPFRFNQVIL-----TAPNRSSVRYDLSLLGPKNFSGILDLDLLLELLE		446
DB	31	FGPYSYTERRIIHDAWPLVDGTFVAGRAVTS--REVQVAVARSGNLLDRVFSNELKL		88
QY	447	LOERLRHLQWSPSAQRNLSLQIDCYAPLNPDTSLYDCCINSLLQYFQNNRTLLLTAN		506

Db	89	MESFIRN-NITVQFSNRTWSFADLCIA--GPDR-----CANNDHIQ-----	127
Qy	507	QTLMGQTSQVDWKDHLFLYCANAPLTFTKOGTALALSCWADYGAPVPFPLAIGGVK-----	560
Db	128	--LASRLHQ-----HGINTYPTVRLSDKSAYIAS-----ALGGVKLAKGDN	167
Qy	561	GKD-YSEAEALIMFTSLNNYP-----AGDPRLAQAKLWEEAFLEEMRAFQRMAGMFOV	613
Db	168	GENIIVENTAWLLIYQLKFYPNEISVVG-----LWEREPKNKMDY-KKQAKYISI	218
Qy	614	TFTAERLSEDEINRTTAEDLPFATSYIVIFLY-----ISLALGSSYSWSRVWVDSKAT	667
Db	219	TYFHSQTLSDLNRAERLAPKEIGAFVILVCFSLCSIVTIKSGYIDW-----VVTXPI	274
Qy	668	LGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRR	727
Db	275	LSVLGVSNAAGMGLASAMGLTYLEIQN-DIIAMFELVAVAGTDNMFVMSASLKRTRD-	332
Qy	728	PGBPREVHIGRALGRVAPSMMLCSLSEACFFLGALTTPMPAVRTFALTSLGLAVIDLPLQ	787
Db	333	-NLKYDQRIAECDADAASVILITALTALSDLSFGVGTITTIIPAVQIFCIYTYMCALLTFAYQ	391
Qy	788	MSAFVALLSDLSKRQE-----ASRLDV-----CCCVKPOELPPGQG-----	824
Db	392	LTFPCAILVYTRIIEEQGLHSIWLRPATVYSSSTPLNKLFWLGSQPK-PLPSCGTVSS	450
Qy	825	-----EGLLGFQKAYAPFLLH-WITRGVVLALLFLALFGVSLYS	863
Db	451	TSSVSTWTSQATSPASKHLHCAATSFRRWYAPVLNQPMI-RAIAGLWLIYIGISYIG	509
Qy	864	MCHISVGLDQELAPKDSYLLDYLFLNRYFEVGAPVFEVTL-----LQYNSFSE	913
Db	510	CTHLKEGLEPANLLVDDSYATPHRVLEKHVHYGASLQIVVSNPPDLRDPVERINMDKM	569
Qy	914	AGMNAICSSA-GCNPFSE-----TQKIQYATE-----FPQSYLAIPASSWDDFI	958
Db	570	ASTFANCKVAIGDSDVSQFWLREMOSVEIHKIQYDNBEKFYDHAQAQYIYSDMSQPVVDVV	629
Qy	959	DWLTSPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLPHFLND	1018
Db	630	-----WGRNN	634
Qy	1019	RENKCPKGLAAVYSTSVNLTSDGQVLASRFMAVHKPKNSQDYTEALRAARELAANITA	1078
Db	635	-----NSERIIKTFRPMIGMRDITSTTTKQTEATNTFREIASRF--	672
Qy	1079	DLRKVGTDPAPFEPYPTITNVFEQYKTLTLPGLFMLSCLVPTFAVSCLLGLDLRS	1138
Db	673	-----EQYNVTYTMPLFLTDDQVALVVPNTMQDIIIVAVACMLVISALLIQQPVCSF	723
Qy	1139	LLNLLSIVMLVDTVGFMAWDISYNVSLINLVSAGMSVEFVSHITRSFAISTKPTWL	1198
Db	724	WV-AVTIGSIDLVGLVGFMTLANNVLDALSMITIMSVGSFSDYSAHITYAVVISKESTTS	782
Qy	1199	ERAKEATTSMGSAFVAGVAMTLPGLVLGLAKAQLIQIFPRFLNLLTLGLLHGLVFL	1258
Db	783	ARVCDALGDLGWPAQAG-AMSTILAVSLSDVPATMI-VTFEFTVFLAISIGFLHGLVFL	840
Qy	1259	PVILS-YVG 1266	
Db	841	PLMLSVFVG 849	
RESULT 13			
T27969			
hypoetical protein ZK675.1 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C:Accession: T27969			
R:Sim, M.			
submitted to the EMBL Data Library, November 1994			
A:Reference number: Z20448			
A:Accession: T27969			

[illegible]

Db 493 WVPFLALGLGVHDHIFMLTAAYASNNR--BQTKL-----ILKVKGPSILFSACTAGSFFA 546

Qy 761 GALTPEPAVTFEALTSGTGLAVILDFLLQMSAFVALLSDSKROEASRLDV--CCVC-----K 814

Db 547 AATPVPALKVFCLOAAIVMCSNLAALIVFPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606

Qy 815 PQELPP-----PGQEGE-----LIGFF 832

Db 607 PKVAPPVPLNNNGRGARHPKSCNNNRVALPAQNPLLEQORADIPGSSHSLSATPA 666

Qy 833 QKAYAPELLHWITRGVVLLIFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLPLNR 892

Db 667 FQHYTPFLMRSWVKFLTMVGLAALITSSLYASTRLQGLDIIILVPKDSNEHFKFLDAQTR 726

Qy 893 YFEVGAPVYFVTTLGYNFSSEAGNAICSSAGCNPFSTOKIOVATFEPQSYLATP--- 949

Db 727 LP--GFYSMTAVTQG-----NFEYPTQQQLRDI--HDSFVRVPHVI 764

Qy 950 -----ASSWDDFDIDLWTFSSCCRLYISGPNKDKFCPSTVNSLNLKNC----- 993

Db 765 KNDNGGLPDPFWLLLFSEWL-----GNLQKIFDEYRDGRLTKECWPNASSDA 812

Qy 994 -----HSITGSRVPSVEQ-----FKYL--PWFLNDRPNIKCPKGG 1028

Db 813 ILAYKLIVQTGHVDNPVDKELVLTNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTL 868

Qy 1029 LAA--YSTSVNLSDGOVLASRFMAVHKLNKSD-----YTEALRAAR 1070

Db 869 LRANCINRANGASQGLYEPQYFHP--NEYDLKIPKSLPLVVAQMPFYLHGLTDTTS 926

Qy 1071 ELAANTPADLRKVPGTDPAPEV--FPYTIITNVEQYLTILPEGLFMLSCLVPTFAVSC 1128

Db 927 QI--KTLLIGHIRDLVKYEGFGLPNYPSGIPFIWEQYMT--LRSSLAMILACVLLAALVLV 984

Qy 1129 LLGLDLRSGLLMLLSVTMLTVDTVGPMALWDISYNVSLINIVSAGVMSVEFVSHITRS 1188

Db 985 SLILLSYMAAVLTVLSYASLAQIFGANTLLGKLSAIPAVILTVSGVMMLCF--NVLTIS 1042

Qy 1189 FAISTKPTWLERAKEATISMGSAVFAGVAMTNPLGILVLGLAKAQLIQIIPFFLNLLITL 1248

Db 1043 LGFMTSVGNRQRRVQLSQMQLSGLPLVHGMULTSGVAVFMLTSPFPEFVIRHFCWLLLVLC 1102

Qy 1249 LGLLHGLVLFVLVILSYVGPVNPALAEQKRABEAAVMVASCPNHPRSVSTADNIYVN 1308

Db 1103 VGACNSLLVPFILLSMVGPEAE--LVPLE-----HPDRISTFSPLPVR 1143

Qy 1309 HS 1310

Db 1144 SS 1145

RESULT 16

T21612

hypothetical protein F31P6.5 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T21612

R:Percy, C.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19449

A:Accession: T21612

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-955 <WIL>

A:Cross-references: UNIPROT:Q19945; UNIPARC:UPI000007B318; EMBL:Z69884; FIDN:CAA93751.1

A:Experimental source: clone F31P6

C:Genetics:

A:Gene: CESP:F31P6.5

A:Map position: X

A:Introns: 64/2; 141/3; 255/3; 308/1; 409/1; 435/3; 484/2; 579/1; 632/2; 715/3; 777/3; 8

Query Match 6.6%; Score 458.5; DB 2; Length 955;

Query Match 6.6†; Score 458.5; DB 2; Length 955;

	Beat Local Similarity	19.8%; Pred. No. 7.7e-24;
	Matches	Conservative 207; Mismatches 402; Indels 227; Gaps 36;
Qy	336	LLGQFFQGMGTWASWPLTLLVLSVPPVVALAAGLV-FTLE-TTDPVELSAPNSQARSE 393
Db	9	LVRNAPWYQVGVHRWRFQFISPLFTLACSGVGFRTMELRUVDDPSYVTPPSDARRRE 68
Qy	394	KAFDQHFQPFRTNOVILTAPNRSSRYRDSLLGPKNFSGILDLDLLELLELLEL-QBRRLR 452
Db	69	ISVFNENW-PL--DENKFLPKGSPEAKRFVNILIRAKDGGSIMEDNVLHEILNQWIMN 125
Qy	453	HLQWSPDAQRNISLQDICVAPLNPDNTSYD--CCINSLLQYFQNNRTLLLLTANQITLM 510
Db	126	NISPTDLDLFNLTQYDCLLS-----YDWVCGANEHIQMLLRNDV-----NOILD 171
Qy	511	QOTSQVDWKDHFLYCAN--PLTFKDGDTALALSMDADYG-APVPFPLAIGGYKGDYSE 566
Db	172	LHFPFGTKDTPVYLGIGDQVQFQNGT-----LSDAKLTQYFLFKQDKQKMWSEYSS 225
Qy	567	AEALIMTFSLNNYPAGDPRLLAQAKLBEAFLEEMRAFORQMAGMFQVTFTAERSLDEIN 626
Db	226	K----PSVAL-----ETFLNQVYS-----SDVITLSFAHYQSLDEGLD 259
Qy	627	RTTAEDLPIDATSYIVIFLYI----SLALGSYSWSRVWVDSKATLGGGVAVVLGAVMAA 683
Db	260	ENAKAFVFNFWSPFLAMTALVSSFTLKSSTATKIDWISKPKWLAAGMFSTVLSTISA 319
Qy	684	MGPFSYIGIRSSILVILQVPPVLV-----SVGADNIFIFVLEYQRLPRP 728
Db	320	FGFLPLUGVRN-VINTIIFLIGEPQSPNMEKQIIPAIGIDDMFMNACWDQTSKSL 378
Qy	729	GEPREVHIGRALGRVAPSMLLCSLSSEACFFLGAALTPMPAVRTPALTSGLAVIDLFLQM 788
Db	379	SVPE--RMSKTLSHAGVAVITNVVDWMSFAICITDLPGLQIFCIYACVSFAFSYFQL 436
Qy	789	SAFVALLSL-----DSKQQR----- 803
Db	437	TFPSGAMAIMGEVEREKRRHCLFFYRTFQLVDISKWNEADSKLQIKRSASPAFPNVL 496
Qy	804	-----ASRLDVC--CVKPOELPPPG-----QSGELLLGFFOKAYAPFL 840
Db	497	SSNSPSSDSDSPSSKTTIPAEFAWKQQQSPNSLSKSKDKREKORIVHFIKGYGFI 556
Qy	841	LHMTTRGVLLLFALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLNRYPEVGAPV 900
Db	557	LSNSVRIFSGLI FVYLAIAWYGCYNFREGNLPCNLVNDHYIAKYFSDIKHFWRIGAQL 616
Qy	901	YFVTTLGYNPSSEKGMNAICSSAGCNFSPQKIQYATBPBQSYLAIPASSWVDDPID 960
Db	617	HVALNPPLNTISENRNEL-----LKVVSAPENTQY-TLREGTGVFFLLEY 661
Qy	961	LTPESSCCRLYISGNPKDKFCBSTVNSLNCNKMSITWGSVRPSVEQFHYKLWFLNDRP 1020
Db	662	LN-----YLSELNAE-----VEDTERLWKTKINSWLKY----- 689
Qy	1021	NIKCPKGLAAYSTSVMLT--SDGOVLASRPMAYHKLKNSQDYTEALRAARELAN-- 1075
Db	690	-----TGSQTQWASNLKINTDGSFQAFRQIALKNPVPENDKHAAQLLADIADHPFN 744
Qy	1076	-----ITADLRKVGVDPAFVPPYITNV-----FYEQYLTLP-----EGLFMLSCL 1120
Db	745	VWVHYEVSFGNRKIL-NDFISSHSCYAQKNIPKLAFFPADQYLTILPATIQNVVISLCLM 803
Qy	1121	VPTFVASCILGLDLRSLGALLSIYMLVDYVGFPMALWDISYNAVSLINLVSAGMSVE 1180
Db	804	-----AVVSFLVPSLPSGFIYFVFSYINSITGVGYMTLWGNVDASWISIMSIGPAVD 859
Qy	1181	FVSHITRSPAIKTPWLERAKATTSMSGSAVFAGVAMTLPGLVILGLAKAQLIQIFFF 1240
Db	860	LSAHIIYAFVTSHGDT-KQRVIGALETLGWPIFOGASST-IAGISILYTDVYII-LVFF 916
Qy	1241	RLMLLITLLGLHLGVLPVILS 1263

Db 917 KTIWLTMLGAIHGLFFPIPIELS 939

RESULT 17

A89153

protein C24B5.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: A89153

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999

A:Accession: A89153

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-956 <STO>

A:Cross-references: UNIPROT:Q18129; UNIPARC:UPI0000076A25; GB:chr_V; PIDN:AAA96120.1; PMID:9851916

C:Genetics:

A:Gene: C24B5.3

A:Map position: 5

Query Match 6.3%; Score 436; DB 2; Length 956;

Best Local Similarity 21.2%; Pred. No. 2.9e-22;

Matches 233; Conservative 177; Mismatches 353; Indels 338; Gaps 48;

QY 345 GTWASWPLTILVSVIPVLAAG-LVFTLETT-DPVELWSAPNSQARSEKAFDQH-- 400

Db 16 GENVHFRFLYLLSFAILLTSAFGFLWFTQTTNDPQVFSPPANAPWRYERAVLTHWP 75

QY 401 -----FGP---FFRTNQVILTPAPNSSVRYDSLLGPNFSGIILDLLELLEQLRL 452

Db 76 LDEKFWPKGSYDLHGVDVIAAGRIHPDY-----GRPN---ILNIRYLDVARINDYII 127

QY 453 HLOVWSPQARNISL-----QDICVAPLNPDTSLVDCINSLLQVFNQNRLLTLLT 504

Db 128 H-----NLTPVDINGKHYDIAT-----DLCMRVDMACYLNDHITMLMP 167

QY 505 ANQ-----TLMGQTSQVDMKDFLYCAN 527

Db 168 KTRWGNFGPPAELASDIINTQVNIPIGWRTETPIYFGALVGAPNLVDEGHFDYASA 227

QY 528 APLTF--KDGTLALSCMADYAGAPVFPFLAIGYKKGK-----DYSEBALIMTSLNRY 579

Db 228 IRLTYNTRGK-----VDQYGT-----WRRKLSMWLTDKENPVSELLEFGVNH- 271

QY 580 PAGDPLAQAKLWEAEAFLEEMRAFQRRMAGMFQVTFPFAERSLEDEINRTTAEGL-PIPAT 638

Db 272 -----NNTLPEGLQD-----VADTLAPKFVG 292

QY 639 SYTVIFLY---ISLALGSYSSRSRVNVD---SKATILGLGGVAVVLGAVNMAAGFFSYLGI 692

Db 293 TCAILTFPSFLSVLRKHRT--GVMPDVRSKPLVAAAGLTPIMATVTSFGLJLWCGF 351

QY 693 RSSLVLQVFPVLVSVGADNIPFVLEYORLPFRPGEPREVHIGRALGRVAPSMLLCSL 752

Db 352 LYN-AIVNVSFPLILCIGIDDLFIMCAEWHR--TNPQHSPEKRIGKTLSEAAVAISITSL 408

QY 753 SEACFPLGALTMPAVRTFALTSLGLAVILDLQ-----MSAFVALL----- 795

Db 409 TDIATFAMGCYTLPFGVQFMCMTVCQCFYVYIIFLGPVLVAYAEHQGHVLLLR 468

QY 796 -----SLDSKRQEAASRLDVCCKVPQELPP----- 820

Db 469 KAVDPDKTESPVKLLLSGSVNRQDQEARRN-----SRKQAVVEKEGKSKFGEVVE 522

QY 821 -----PGQG--EGLLIGFTQKAYAPPELLHWITRGVVLLIFLALFGVLSYMSCH 866

Db 523 KLEHTLEKHHDDPGHNSSEETLVSKVFRFIIIGPFILOKSTQVCALLLYLVYISAGGCLN 582

QY 867 ISVGLDQELALPKDSVLLDYFLFLNR-YPEVGAPVYFV-----TTLGYNFSSEA 914

Db 583 IKEGLDPKLLVRESFYLSKFVEIIDETFWREGLQVQVWVNNPPDLFTPTETRKQFD----- 637

QY 915 GMNAICSSAGCNPFSTQKIQTATEFPEQSYLAIPASSWVDDDFIDWLTPTSCCRLYISGP 974

Db 638 -----EMMAEPENTQYTHPNATMI-----WL-----RAY----- 662

QY 975 NKDKFCPSTVNSLNCNCKSIITWGSVRPSVBOPHKYLPLWFLNDRPNIKCPKGGLLAYST 1034

Db 663 --EHLLETEVHELNIK-----PNSVTEWTRCDWLI-----VAGGRLLWQM 703

QY 1035 SV---NLTSQ--GQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVGTDPA 1089

Db 704 DMVWANTTETPRITAFRQGLRNYRTPTDHTHSCKLMBSIA-----DK 749

QY 1090 FEVPFYITNVFY---BOYLTILPEGLP--MLS-LCLVFPFAVSCLLGLDLRSGLLNLL 1143

Db 750 YSMFNVTTFHEYYPFADQYLELTP-SLFQNMISDLCTI--FAVSWVMIP-ELICAVAIVL 805

QY 1144 SIVMILVDYVGFPMALWDISYNAVSLINLVSAVGSVFEVSHITRSFAISTKPTWLERAKE 1203

Db 806 SIASINVGVLGFMFGVGNLDSVSIITVICIGFSVDLSAHIAVAFSQSYGNSH-ARAVA 864

QY 1204 ATISGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFRLNLLITLGLLHGLHFLVPLVLS 1263

Db 865 ALETLGMPVPLGASSTVL-GILLTLVDSYIVQIFFKTVFLVIN-FSILHGLIFLPIFLM 922

QY 1264 YVGDPVNPALALEOKRAEEAV 1284

Db 923 KVVREVK-----ESPRDDQSV 939

RESULT 18

T26521

hypothetical protein Y18D10A.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T26521

R:Harris, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: Z20226

A:Accession: T26521

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-936 <WII>

A:Cross-references: UNIPROT:Q9XW22; UNIPARC:UPI00000813DB; EMBL:AL034393; PIDN:CAA22312

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.7

A:Introns: 11/1, 44/2, 106/1, 166/3, 289/2, 373/1, 396/3, 485/1, 573/3, 615/3, 644/3, 651/3

Query Match 5.9%; Score 409.5; DB 2; Length 936;

Best Local Similarity 20.2%; Pred. No. 2e-20;

Matches 213; Conservative 172; Mismatches 369; Indels 301; Gaps 40;

QY 342 CQSGT-----WVASWPLTILVSVIPVLAAGLVFTLTDVPVELWSAP-NSQARSE 393

Db 29 QPWANIVAKYCLFVAKYWPFFIILITICLSMGIILNFKIVRGVNYLYAPLNATWKE 88

QY 394 KAF-----HDQHPGPPFRTNQVILTAPNRSYRYS--LLGPNFSGIILDLLEL 444

Db 89 EAVFGENWAKODDF-----YPGKDIRRQGIYLVNNAKGGNVLNRQEAQDF 136

QY 445 LEQLERHLQVWSEARQNSISLODICVAPLNPDTSLVDCINSLLQVFNQNRLLTLLT 504

Db 137 LKILDMILNKLSS-SAGRIFTYKQVC---LHFQN-----DC-----FSPHAKLL-- 178

QY 505 ANQTLMGQTSQVDMKDFLYCANAPLTFKDGTLALSCMADYAGAPVFPFLAIGYKGY 564

Db 179 -----ANIYSKHQNSMENIYPIVRS-----YATEPIDISKVLGNVSLDY 220

QY 565 -----SEAEALIMTFSLNN-----YPAGDPLRAQAKLW---BEAFLEEMR 601

Db 221 DGHVENASAWLILYQLKNEKWLSDRDFEDGLAEKIQSGEAPSEBLLNLYYFHSATPDQBLE 280
QY 602 AFQRMAGMFQVFTTAERSLEDEINRTTAEDLPFA--TSYIVIFLYISALG-----SYS 655
Db 281 KENRRUTPKFSITFSV-----LIIFAITTFTIKPMKPTENGINGQYPI 325
QY 656 SHSRVMVDSKATILGLGVAIVLAVMAAMGPFSSYLGRSSVLQVVPFLVLSVGDNIIF 715
Db 326 DM-----VLSKPLIGICGLVLTWCATIISSGLLMLFN--TFVDMCTVMPFLSTIGIDDTF 380
QY 716 IFVLEYORLPFRPGEPREVHIGRALGRVAPSMCLCSLSAICFFLGLATPMPAVRFTALT 775
Db 381 LMLAAWHETR--NLPEYKRIEKMRHAAVSIISLTDLALFLIGSIAPLPAVIYFCY 438
QY 776 SGLAVILDFLQMSAFVALLSDSKQASRLDVC-----PFOKAYAPFLHWTGCVLLFLALP 857
Db 439 SAAAILFELYVLTMEVAVLALQGREEDLKHSVTGMKTIDLDSDYETASTRQLLLKMS 498
QY 811 -CCVKPQELPPQGGGLLG-----PFOKAYAPFLHWTGCVLLFLALP 857
Db 499 RVSVKADEENNNNEKSIENIKIDNRMYORFFEDQYAPFISNKISILSFLIYLAFL 558
QY 858 GVSLSYMSCHISVGLD-----QELALPKDSYLLD 885
Db 559 AAAYGVKRLKIGDFVTVVLTIKYKVLQINIVQEDSASRVFLEVRQLPEDTKLMD 618
QY 886 YFLFLNRYEPVAGPVYFVTTLGYNFSSEAGMNAICSSAGCNFSPFKIQYATEPPEQSY 945
Db 619 IAV-----MNSP-----NFSNPE-----ERNFP-----NEVLSEP----- 643
QY 946 LAIPASSWDDPID-----WLTSSCCRLYISGPNKDKFCPTVNSLNCUKNCSITMGSV 1001
Db 644 ----ESTWCSEGRESTQFWFFE--MQKYLNSLFGGDLTKWNS-----ERKLSQSKTF 692
QY 1002 RPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSNLTSQGV--LASRPMAYHKLKNSQ 1060
Db 693 LMSHEKF-----GVDVSDKQFRLSTR-----LKNVE 719
QY 1061 DYTEALRAARELAANITADLRKVPDPAFEPFYTITNVFQYLTILPEGLFMLSCL 1120
Db 720 TDEMFENCAR-----TWKLSQKHANYSIITVSLPWNIADEYDIMPQTMQDIYISI 771
QY 1121 VPTFAVSCLLGLDRLSGLLNLSIVMLIVDTVGFMAWLDISYNAVSLINLSVAGMSVE 1180
Db 772 AVMPVALLFIQPLCSVIIG-LNITASIAFGVIGTMSFLGSLDATSMITVAMSFGSVD 830
QY 1181 FVSHITRSPAISTKPTWLERAKENTISMGSAVFA-----GVAMTWLPGLVLGLAK 1231
Db 831 FAHV--SYAYMTE-----SGAQIPGKSATYSRPFCHTLGTIGTPVTOASVSVLLGVSS 881
QY 1232 AOLIQIFF-----FRLNLLITLLGLHLGLVFLPVIL 1262
Db 882 LVLVDSYVQTCFRTVVLVILFGTHALVFLPULL 916

RESULT 19
T26683
hypothetical protein Y38F1A.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26683
R:Wallis, J.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20253
A:Accession: T26683
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-800 <WIL>
A:Cross-references: UNIPROT:O9KWL9; UNIPARC:UPI000007721A; EMBL:AL032639; PIDN:CA021636
A:Experimental source: Clone Y38F1A
C:Genetic:
A:Gene: CESP:Y38F1A.3
A:Map position: 2

A;Introns: 13/2; 100/2; 132/3; 220/3; 257/2; 521/2; 642/3; 710/3
Query Match 5.9%; Score 407; DB 2; Length 800;
Best Local Similarity 23.4%; Pred. No. 2.5e-20;
Matches 175; Conservative 119; Mismatches 285; Indels 168; Gaps 27;
QY 594 EAFLEMRAPQRMAG-----MFQVTTFAERSLEDEINRTTAEDLPFAFATSIIVIF 644
Db 121 ETMYIMKEWEOKLPAYTLSTENDPLIRVYVTTSEGLVSEVRRTGILAMPLMGVTFLLI- 179
QY 645 LVISALGSYSSWSMVDSKATILGLGVAIVLAVMAAMGPFSSYLGRSSVLQVVPFL 704
Db 180 ----LAFITLITLKDKDPVKSKEPFAFLVVICILCASFGHLFWMGF-EYPIVITVVP 234
QY 705 LVLSVGADNIFIFVLEYORLPFRPGEPREVHIGRALGRVAPSMCLCSLSAICFFLGLAT 764
Db 235 LILSIGVDDVFIHAWHRTPYK--HSVRDRMAETLADAGPSISITSLNLSFAIGIT 292
QY 765 PMPAVRTPALTSGLAVIDFLQMSAFVALLSDSKQASRLDVCCKVKPQELP--PPG 822
Db 293 PTPAIYVTCVFISTAVIYDIYIYQIFPFSAVLVLSGEREE-QRKNAYLMMKDVLPPEPTG 351
QY 823 QGE-----GLLIG-----PFOKAYAPFL--HWI--TRGVVLLLLFLALPGVSL 861
Db 352 KQKKEIVSRVLGKVDLFWVDFIMAKWSKFLIGAIMITYMIFMARGVM----- 399
QY 862 YSMCHISVGLDQELALPKDSYLLDYF-LFLNRYEPVAGPV-YFVTTLGYNFSSEAGMNAI 919
Db 400 ----EIAVGLSEKFLDSDSPLLPLVRQTNTVIFKEGGQVAVFVNNPG-NMSEPDVAP 454
QY 920 CSSAGCNFSPFKIQYATEPPEQSYLAIPASSWDDPIDWLTSPSSCCRLYISGPNKDK 979
Db 455 -----MRILRRFEVANNVSGAASTHMWLLPYLPYVGEQ- 488
QY 980 CPSTVNSLNCUKNCSITMGSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSNLTS 1039
Db 489 -----HGSIEFKYRILPEPFK-----LTFEFRMWSHFVNLG 518
QY 1040 SDGQVLASR-----FMAYHKPLKNSQDYTEALRAARELAA-----NITADLRKVP 1085
Db 519 NHQDCLSEKPSCLKQFVSTGFHDVSW-S-DRLALLENWREMASEYQHLNLT- 570
QY 1086 TDAPEVFPYTTITNVFQYLTILP---EGLFMLSCLVPTTFAVSCLLGLDRLSGLLN 1142
Db 571 -----YEDFSMYSQDLITIVPTQSTVFCALICMI-----MIUTLTPSPVTV 614
QY 1143 LSIWML---VDTVGFMAWLDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLE 1199
Db 615 TSTAVALSINLGVFCGLVYMNIDLDPIGMTTLMAIGSFSDVFAHITWHYYKGFHSHKRA 674
QY 1200 RAKEATISMGSAVPAVAVMTNLPGLVLGLAKAQIQLQIPFPFRLNLLITLLGLHLGLV 1259
Db 675 RIRHALAGIAMPWFOAGTSTML-AITVLALVHAYMVQV-FVKVWVVLVIFLGMFHGLVLP 732
QY 1260 VILSYV-----GPDVNPALAEQKRA 1280
Db 733 IVFSALPFTKTSQPKKKVAPLQMEHA 759

RESULT 20
T31739
hypothetical protein T21H3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31739
R:Scheet, P.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid T21H3.
A:Reference number: Z21077
A:Accession: T31739
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-881 <SCH>

Db 628 TFAFKISNNIRTSLLLTWRNTSSYPEDDALVFDENNFSYDQMBLQSTISSLGTALII 687
QY 1115 MLS-LCLV-----PTFAVSCILLGLDLASGLINLLISVMILVDTVGFMAWDISYNAYS 1167
Db 688 LLSAICIMPIAESYIVFVVVCTIISND-----IGTAGLLSLMGADLDPM 732
QY 1168 LINLVSAGMSVEFVSHITRSPFAISTKPTWLBRKATISMG-SAVFAGVAMTNLPGLIV 1226
Db 733 VVNILMSIGQCIDFATHGIRIYRSYSDPDERIDAMGAIGWPPVQAGT--STLLAIVV 790
QY 1227 LGLAKAQLQIPFFRLNLALITLLGLLHGLVFLPVIL 1262
Db 791 MLMVPSSARM-FARTSVILVVGTFPHGLIILPIIV 825

RESULT 22
T16126
hypothetical protein F21H12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16126
R:Pavello, T.
A:Description: The sequence of C. elegans cosmid F21H12.
A:Reference number: Z18464
A:Accession: T16126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-714 <PAV>
A:Cross-references: UNIPROT:Q09540; UNIPARC:UPI0000076F08; EMBL:U23176; NID:G726404; PID
A:Experimental source: Strain Bristol N2
C:Genetics:
A:Gene: CESP:F21H12.4
A:Introns: 25/1; 334/2; 430/2; 543/3

Query Match 5.5%; Score 379; DB 2; Length 714;
Best Local Similarity 23.0%; Pred. No. 2e-18;
Matches 173; Conservative 119; Mismatches 299; Indels 160; Gaps 25;

QY 545 DYGAVPPLATGGYKGDYSEBALIMTFSLN--NYP-----AGDPLAQ--AKLWERAP 596
Db 58 DYAEW-----TGCGSG-----FASVLNWPEDMILGNPRAKGGKLSGADA 100
QY 597 LEEMRAFQRRMAGMFOVTFABRSLEDEINRTTAEDELPIFATSYIVIFLY----- 646
Db 101 LQ-----FVPLVASPADVFLRFKQSLER-----TLETLSLYEHCHLHRHPLICSRNSV 151
QY 647 -----ISLALGSYSRSRVNV-----DSKATILGGVAVVLGAVMAAMGFFSYL 690
Db 152 NSTIPSSLLDMLSCWHMLLLILLDLTTAPRLQHQWGFALGVVTFASVAGIGLATWF 211
QY 691 GIRSSVILQVPELVLVSGADNIFIVLEYQRLPRRPGEPREVHIGRALGVAPSMLLC 750
Db 212 GLEFNAATQIVFFETLIGVGNMFMNLHNYDDVKLAGHAEMAI--LMRETGMSILCT 269
QY 751 SLSEAIICFPLGALTMPAVRTFALTSGLAVIDLLOMSAFVALLSLSKROEASRLDVC 810
Db 270 SINNLISFLTGTLLPILALRSFCAQSSILLTFNTFALTITPAIISIDLRKKAQRDPD 329
QY 811 CCVK-PQELPPPPQEGEL-----LLGFFQKAYAPFLHWTITRGVLLLLPL 854
Db 330 TREESTSEASIMQFDGITOQAQWASSDDDPAPWSLSHFIRYIYIPFISKPAKVAITVGCC 389
QY 855 ALPGVSLYSMCHISVGLDDELALPKDSYLLDYFLFNRYFEGAPVYFVTTIGYFNS---- 911
Db 390 ALLGASFVGMRQSTLGLGLDGLVLPHTAPAQFLRADKDYFSP-YPM-FAVIRKPNIDYAH 447
QY 912 SEAGMAICSSAGCNNSFTQKIQVATPEEQSYLAIPASSWVDDFDIMLTPSSCRLYI 971
Db 448 QQRQIDNTYKQISGSSKIYVTKN-----NEEPSEKY-----WGLMRDWLI--STORGPD 494
QY 972 SGNPKDKF-----CPSTVNSLNCNCSITMGSVRPSVQFHK--YLPWFLNDRPNI 1022

Db 495 BEVAGSGFNLTSGTIVGSNVSEDAHLAALMCSHGLSPLFCAGRIGKIRLIPFYL----- 549
QY 1023 KCPKGLAAYSVSNLTSDQVLASRFEMAYHKPLKNSQDYTEALRAARELANITADLRK 1082
Db 550 -----GLTDTAVIVDAIKDIRSVCRF----- 571
QY 1083 VFGTDPAPFVPTTNTNPFYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRSGLLNL 1142
Db 572 ---TQOGLPNFPQGIAPTFWEQYLFITGNLMQAISIITISVFCVISVLL-FNPWAAALMVV 627
QY 1143 LSIIVMILVDTVGFMAWDISYNASVLINLVSAGMSVEFVSHITRSPFAISTKPTWLBRK 1202
Db 628 CILGIMTCELAGFMGLVGIKLPNSAVTLITAVGIGVEFTVHVVSFL-----TALGTRS 682
QY 1203 EATISMGSAVFAGV---AMTNLPGILVLGLA 1230
Db 683 QRTSSAVDRVFVPIVHGSFSTLLGILMLGFS 713

RESULT 23
T23399
hypothetical protein K07C10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23399
R:Coles, L.
A:Submitted to the EMBL Data Library, December 1994
A:Reference number: Z19736
A:Accession: T23399
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-690 <WIL>
A:Cross-references: UNIPROT:Q09934; UNIPARC:UPI00000835CB; EMBL:Z47074; PIDN:CAA87375.1
A:Experimental source: clone K07C10
C:Genetics:
A:Gene: CESP:K07C10.1
A:Map position: 2
A:Introns: 34/3; 362/3; 448/2; 507/3; 548/1; 614/3

Query Match 5.3%; Score 367; DB 2; Length 690;
Best Local Similarity 21.3%; Pred. No. 1.3e-17;
Matches 163; Conservative 148; Mismatches 314; Indels 140; Gaps 27;

QY 564 YSEBALIMTFSLN--NYP-----PAGDPLAQ--AKLW-----BEAFLEEMRAFQRRMA 608
Db 8 YPLAEAFGHFPYGLNQFPFGVNGQASKAGPIKTAKEVALWYMSKAETP-EQOKQLQAVQL 66
QY 609 GMFOVTTAERSLE-----DEINRTTAEDELPIFATSYIVIFLYISLALGSYS 656
Db 67 GIFKKSASRPNNLQPDFEMFGDOVANSEMLRGTLTTVKLFFIGGCLMWAFMACTFTELTI 126
QY 657 WSRVWVDSKATILGGVAVVLGAVMAAMGFFSYLGIKRSVLVQVVPFLVLSVGADNIFI 716
Db 127 FSKOMLIAGA---IGSPLAATGACPAIILGWVGH-----PFNSIMCITPFLILGIGVDDAFL 179
QY 717 FVLEYQRL--PRRPGEPREVHIGRALGVAPSMLLCSLSSEAICFFLGALTMPAVRTPAL 774
Db 180 LINCNRREGKDKSAKQAEQLARVIREISPMALTSITNTNWFVGLFAPLTPQMSFCL 239
QY 775 TSGLAVIDLLOMSAFVALLSLSKROE---ASRLDYCCCVKQPELPPPGQEGELLGF 831
Db 240 GTALAIVLDELLEFLIFVPCWVLFYKKKEPIIAIENEEFANEKTEKRSSTSTGKYSRINW 299
QY 832 FQKAYAPFLHWTITRGVLLLLPLAFGVSLYSMCHISVGLDDELALPKDSYLLDYFLFN 891
Db 300 --KCFNTLLSVPGRVLVILYFSTASYFGVAKMEKTFDPSKTFPSPSKLVDLSASF 357
QY 892 RYFVEGAPVYFVTTIG--YNFSSEAGMAICSSAGCNNSFTQKIQVATPEEQSYLAIP 949
Db 358 SIQVYSPLNFLSSVFDLKNDDTDVANFN-----EMLHKLBFREECYGDVG 402
QY 950 ASSWVDDFDIMLTPSSCRLYISGPNKDKFCPSTVNSL-NCLKNCSITMGSVRPSVEQF 1008


```
QY 620 SLEDEINRTAEDLPIFATSYIVIFLYISIALGSSYSWSRVMDVSKATLGLGVAVVLGA 679
Db 188 VLDNEMKQGRRLTPFAAGFOFEITFVSLCVILTITAVYHNCIDQGGKLLISLGLAVLCPILA 247
QY 680 VMAAMGFFSYLGRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHVHIGRA 739
Db 248 ISTYGIIVILGNRTNSFML-VMPFLVMGIGVDSCLFMTHSQKERRAQAASGTGNRLGMV 306
QY 740 LGRVAPSLCSLSEAICFFLPGALTPMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDS 799
Db 307 YESVSGSITITSLTDFLSFAIGALAPTETRLFCIASIALALTYILQLVLFGLPILAVAT 366
QY 800 KQCEASRLDVC---CVKPOBELPPQCGEGILLGFQ-----KAYAPFLHMITRGV 848
Db 367 RYEHKTTTTNCRNWKWVRIR-----LGDGIIFSKLKAKCKKCNVYCKILSNKPPAVV 420
QY 849 VLLFLALFVSLYSCHISVGLDQELALPKDSYLLDYFLNRYEPCVAPVVFVTTLGY 908
Db 421 VMLGTAIYVFAIYGLMTWKTRLDVAKIILPKDSPLQRPNLVLTNL--VWANTHPVTIL-- 476
QY 909 NFSSEAGMNAICSSAGCNFFSTQKIYATEPPEQSYLEAPAS--SWVDDFDLWLTTPSSC 966
Db 477 ---INAPLDLE--NRHOMDRYWNMVDPEPEKHCHKGKASTLSWLRLDYIKFSYHGEP 527
QY 967 CRLYISGPNKDFECPSTVNSLNCNMSITGWSVRPSV-BOFHKY-----LPWFLN 1017
Db 528 FNLF-----AFFGLMTPPEVYEEVDPPYQANITTAFLPEFLK 562
QY 1018 DRPNIKCPKGLAAYSTVNL-TSDGOVLASRFMAVHKPLKNSQDYTEALRAARELAANI 1076
Db 563 S-PFFK-----HWSFIHYHEGVIKLDRFM-----MNWAYONT---SSWDTRIQL 605
QY 1077 TADLRKVPCTDPAFEPFPTIITNVFEQYL-----TILPEGLFMLSCLVPTFAVSCLLIG 1132
Db 606 MTDWRKVANNYSDLNVTVWEPNGMFDQMLSLGRATATQGIWTL-VCM---AVVCAIFI 660
QY 1133 LDLRSLGLLNLISVIMLVDTGFMALWDISYNAVSLINLVSAVMSVFPVSHITRSFALS 1192
Db 661 PNPCSIITATSVIASITITGVMGFLSLWSPDLDPVVMAAVLMISGLSVDFIAHVAYHFOLA 720
QY 1193 TKPTWLERAKEATISM-GSA-----VPAGVAMTNLPGLIILVLGLAKAQLIQI 1237
Db 721 HRKE-IRNGKIKKIPUGSTERLEHTLGAVAMPQAGV--STICCILPL-LFRASYSPS 776
QY 1238 PFRLMLLITLGLHLGLVFLPVILSYVGPVNPA 1272
Db 777 VFVAFLVTVTGMLHGLLILPTFLAALPESVTTA 811

RESULT 26
T32908
hypothetical protein F56C11.2 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C.Accession: T32908
R.Tin-Wollam, A.; Wohldmann, P.; Morris, M.
submitted to the EMBL Data Library, January 1998
A.Description: The sequence of C. elegans cosmid F56C11.
A.Reference number: Z21244
A.Accession: T32908
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-820 <TIN>
A.Cross-references: UNIPROT:044978; UNIPARC:UPI000007C788; EMBL:AF043697; PIDN:AAB97556.
A.Experimental source: strain Bristol N2; clone F56C11
C.Genetics:
A:Gene: CESP:F56C11.2
A:Map position: 1
A.Introns: 24/1; 58/2; 97/3; 134/3; 171/1; 200/3; 332/2; 368/3; 479/2; 503/3; 580/2; 678/2

Query Match 5.0%; Score 347.5; DB 2; Length 820;
Best Local Similarity 20.5%; Pred. No. 3.8e-16;
Matches 198; Conservative 161; Mismatches 390; Indels 219; Gaps 41;
```

RESULT 27

T22186

hypothetical protein F44F4.4 - Caenorhabditis elegans

```
QY 343 GW-----GTWVAGWPLTILVSVIPVVALAAGLVFTLTTPVWELWGSAPSQARSEKA 395
Db 2 GMDVISAVVARLAAPVYSILLSILSVLSGLVHRLPEDIRKAFSPDDSDAGYETR 61
QY 396 FHDQHFPGPPFRNOVILTAPNRSSRYDSLLGPKNPFSGIILDDLLLELLELQERLHQ 455
Db 62 VWLEYG-----LDIYPERAF-----CIPTAKRENSILOEGALADITYVDNRL----- 105
QY 456 VMSPEAQRNISLQDICYAPLNPDNTSLYD--CCINSLLQYFONNRTLLLTANQTLMGOT 513
Db 106 -----SAAVGQD-----GDGRKNCPLCDLMSPPHLLAN-----LTRNN--GTT 144
QY 514 SQVDWKDHFLYCANAPLTPKD--GTALALSCMADYGAPVFPFLAIGYGKGYSEABALIM 572
Db 145 SVFTYPD-----MPYSGLDIFLGLHLS--NADFPPT-----ERISSKSLVLM 185
QY 573 TFSLNYPAGDRLPAQAOKLWEAFLEEMRAFORRMAGMPOVTPT--AESRLEDEINRTTA 630
Db 186 YFSRSDTPEG-----KLAFKDAIDELFQLSKSSAFEHVQVTFISDQVANREMRIGAI 238
QY 631 EDLPFATSYIVIFLYISIALGSSYSWSRVMDVSKATLGLGVAVVVLGAVMAAMGPFSL 690
Db 239 EATLTMTIGFFLLLTQVILVIRLS-----VOMAFYLVAATSLTPMAATVASFGAICWM 293
QY 691 GIRSSLVILQVVPFLVLSVGADNIFIFVLEYQ-----LPRRPGEPREHVHIGRALGRVA 744
Db 294 GP-PSFSIQCVTPFLVIGVGDDAFILLHRWKHHTAITDTPR-----LEQVIVDVG 344
QY 745 PMLLCSLSSEAICFFLPGALTPMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSL----- 797
Db 345 PSITITSLNITIAFGIGELFTPTPQMSLFCLTASLALLDYIFTYITILAPIVELCNDPDYK 404
QY 798 ---DSKROBASRLDVCCKVQBELPPQCGEGLLGLFPQKAYAPFLHMITRGVLLFL 854
Db 405 STANTEPEKISRADT-----WLSKYSSFCVSKVRACIMVLL 442
QY 855 ALFGVSLYSCHISVGLDQELALPKDSYLLDYFL-----PLNRYEPCVAPVVFVTTLGYNF 910
Db 443 GMYALTITGVVMTKTTTPPAKAFPSNKLVDLSNIPKVFNYF---PIITVINPPNI 498
QY 911 SSEAGMNAICSSAGCNFFSTQKIYATEPPEQSYLEAPASVWDDFDLWLTTPSSCCRLY 970
Db 499 EIKAEYDSF-----NN--MMDRLEH-----VPG----- 519
QY 971 ISGPNKDK-FCPSTVNSLNCNMSITWG-SVRPSVQFHKYLP-WF--LNDRNKICP 1025
Db 520 IRGDNRSILIFQYVN-FDKTNMILSTLVGEKIRPSYN-----LPTWMDAIGNPPLVKYH 574
QY 1026 KGGLAAYSTVNLTSDGQVLASRFMAVHKPL--KNSQDYTEALRAARELAANITADLRKV 1083
Db 575 MG-----DDNKTIVT---APRLTLGKMSWAERARAKQH-----IRTI 611
QY 1084 PGTDPAFVFPFPTIITNVFYQVLTILPGLFMLSCLVPTFAVSCILLGLDRLSGLNL 1143
Db 612 LHEESQFNATLPDCDSAILSITLTGTDLIGSIATVW-CMAIVCFVFTANFNAVAVITS 670
QY 1144 STVMILVDTGFMALWDISYNAVSLINLVSAVMSVFPVSHITRSFALSTKPTWLERAKE 1203
Db 671 VIASICYVLVGLSLWGADLPVIOVDVLLATGFSVDYTAHVAYNF-FRARTGPQERVYS 729
QY 1204 ATISMGSAVF-AGVA--MTNLPGLIIVLGLAKAQLIQIPFRLNMLLITLGLHLGLVFLPV 1260
Db 730 SLAEMAMPCEAGLSTFLCMLPLIFVPIYA-----IVCFAKTVFLVAIGLHGLFLEPV 784
QY 1261 ILSVVGPD 1268
Db 785 ILALPSKD 792
```

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22186
R;Colter, L.
submitted to the EMBL Data Library, September 1994
A;Reference number: Z19528
A;Accession: T22186
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-890 <WIL>
A;Cross-references: UNIPROT:Q20404; UNIPARC:UPI0000061044; EMBL:Z37092; PIDN:CAA85455.1;
A;Experimental source: clone F44F4
C;Genetics:
A;Gene: CESP:F44F4.4
A;Map position: 2
A;Introns: 51/2; 148/3; 183/3; 254/1; 374/3; 573/2; 623/2; 665/2; 683/3; 769/1

Query Match 4.5%; Score 312; DB 2; Length 890;
Best Local Similarity 18.7%; Pred. No. 1.3e-13;
Matches 188; Conservative 181; Mismatches 396; Indels 242; Gaps 39;

QY 345 GTWASWPLTILVSVIPVVALAAGLVFTELTDPVELWSAPNSQARSE-KAFHDQHP-- 401
DB 22 GSTVARYPVITIFCLVLSAVMSIKALTP-KVSAIDGVASDDSMRHEFKTF--QQFLD 78
QY 402 --GPFRTNQVILTAPNRSSRYSDLSLLGPKNFSGILDLDLLELLEQERL---HLQV 456
DB 79 SDGPGITTAILLI-----RSSITNGSL-----LBEQRLKEVVKVDSFISTNFKLV 123
QY 457 WSPQAQRNLSQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMOQTSQV 516
DB 124 SGVEKNFQFCRGFOA-----NEPRVQYNG--LQILGNQTF-DQISKRI 166
QY 517 DWKDHFLYCANAPLTPFKOGTALALSCMADYGAPVFPFLAIGYKGYKDYSEAEALIMFSL 576
DB 167 D-----LSVPTSNFFGKESL-IPNFFGLSMSP-----EQHLNSSLNLIIVLYFRA 210
QY 577 NNYPAGDPPLAQAQKWE-----BAFLEEMKAFORMMAGMPOVTFTAERSLDEBINTTAD 632
DB 211 ERYPP--DWSTKTQKQWELRVDRHFAKEY-----SSDLIIIVDMSTQTVVESEIVRAGLSL 262
QY 633 LPIFATSYIVFLYISALGSSYSSSRV--WVDSKATLGLGGVAV--LCGVNA-----AMGF 686
DB 263 QPELIVGVFVMSIFCTVT---TWFSAVLYISQKATFNKVALSIACINPFWACGTAGI 318
QY 687 FSYLGIRSSIVILQVVPFLVSVGADNIFIFVLEYQRLPRPCEP-----REVHIGRALG 741
DB 319 LFFCGVTFS-PINCIITPFLVLAISVDLSFLMLHAWNRLSRTAPLDPKMRHMMGEVLV 377
QY 742 RVAPSMLLCSLSRAICFFLGALTMPMPAVRTFALTSGVLVDLLOMSAFVALLSI--DS 799
DB 378 ETGPAISISAFNTMLAFTIGAITSPPEIRIFCFGNAAIIFMDMFYQATFYTACMTLLGT 437
QY 800 KROEASRLDVCCCKQPELPPCGGGLLPFOKAYAPFLHWTIRGVVLLLFALFGV 859
DB 438 KNEG-----VSEKTRIQ---ERGNVVGRFLKNYVNAISNI FVSFGIVLWAVFIGF 488
QY 860 SLYSMCHISVGLDQELAKPSYLLDYF--LFLNRYFEVGAPVYFTTLYGYNFSSBAGM- 916
DB 489 AVLGLTRHLVLEPSPKFFLKDSPML-YMDRLRTNEVVVYTPVHVIVNPNPGDLTNDNSVE 547
QY 917 -----NAICSSAGCNFSFTQKIQVATEFPESQSYLAI-----PASSWVDDF 957
DB 548 RLVELKEKLEHMPNAI--GAPSTKFFLDLPVQYRSSFAEEIEMDIADDESEKSDIEQF 605
QY 958 IDMLTPSSCCRLYISGNPKDFCPSTVNSLNCNKMSITMGSRVSPVQFHYLPWFNL 1017
DB 606 LEW-----PEFSFWRGFLAFDQ 623
QY 1018 DRPNIKCPKGLAAYSTSVNLTDGQVLASRFM---AYH-KPLKNSQDYTEALRAARELA 1073
DB 624 THPH-----NVT-----KFMFTTGFGQDLKDWKRGQLLRWGA 660

QY 1074 ANITADLRKVPGTDPAPFEVFPVTITNVFYEQVLTILPEGLFMLSLCLVPTFAVSCLLGL 1133
DB 661 BEFKDDFNTVTFEDA-----FFLDMDISPTTWTQTLATFFVSLVCLFIS 709
QY 1134 DLASGLINLLSIWMLVDTVGFMALWDISYNAVSLINLNAVSGMSVEFVSHITRSP---- 1189
DB 710 DILTVMLSVATLVTSGVFGYLSLLGVTLDPVMSIATMCIIGSVSDIPARVAFHFYAAK 769
QY 1190 -----AISTKPTWLERAKENTISMGSAVPAGV 1216
DB 770 AKKHASSHSQPSHQVTSPTDIVHSTSTSSDSSISPTNPSFETNLPALSSVGPVVIQAG 829
QY 1217 AMTNLPGLVLGLAKAQIQLIPFFRNLTLITLGLHGLVFLFVILS 1263
DB 830 VSTDPCA-LPLGFMELYMAKMFALSITLQVLS-LSLHGLIVIPALLS 874

RESULT 28
T28706
hypothetical protein C41D7.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T28706
R;Rohlfing, T.; Wohldmann, P.; Le, M.
submitted to the EMBL Data Library, August 1997
A;Description: The sequence of C. elegans cosmid C41D7.
A;Reference number: Z20513
A;Accession: T28706
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-871 <ROH>
A;Cross-references: UNIPROT:O16530; UNIPARC:UPI000017B827; EMBL:AF016653; PIDN:AAC48225.
A;Experimental source: strain Bristol N2; clone C41D7
C;Genetics:
A;Gene: CESP:C41D7.2
A;Map position: 2
A;Introns: 34/3; 110/3; 173/3; 216/2; 264/3; 331/2; 394/3; 501/3; 625/3; 818/3

Query Match 4.3%; Score 297.5; DB 2; Length 871;
Best Local Similarity 20.0%; Pred. No. 1.3e-12;
Matches 202; Conservative 168; Mismatches 432; Indels 207; Gaps 41;

QY 353 LTIILVSVIPVVALAAGLVFTELTDPVELWSAPNSQARSE---KAFHDQHPG---F 404
DB 2 LTLVLSLKIP-----FTRMQND-ISDTPFGARSVELEKYRKFNNHGPKAIFY 51
QY 405 FRTNQV-----ILTAPNRSSRYSDLSLLGPKNFSGILDLDLLELLELQ 449
DB 52 FITSKVGDNMLGISQLNDTVHVLDTISRDFYLKTS--QGPKNFE-----EYCSGFCLLNE 104
QY 450 RLRLH---QVWSPQAQRNLSQDICYAPLNPDNTSLY-DCCINSLLQYFQNNRTLLTLTA 505
DB 105 PVKHPYSKMLISGQGTDFSHLDLGYPTITVLTGLYMDPNFPGVKVTTSSNNKVESIADG 164
QY 506 NQTLMGQTSQVDWKDHFLYCANAPLTPFKOGTALALSCMADYGAPVFPFLAIGYKGYKDY 565
DB 165 NTKLFG-----DH---TKQEPNNIREVSLVVLQFSELGDEVTP-----EDLR 204
QY 566 EBAALIMFTSLNNYPAGDPRLAQAQKLEEAFLSEBRARQRRMAGMFOVTTTAERSLDEBI 625
DB 205 NYEQILDYIHDEY-----QSEHINVIILTDSTYITEI 237
QY 626 NRTTAEDLPFATSYIVIFLYISALGSSYSSSRVWVDSKATLGLGGVAVVGLGVAAMAAG 685
DB 238 VRAGLTLLPFLVIGFTIMAVFSSI-----TFVPIAWMACVCPFMACGSLGAMF 287
QY 686 PFYSGIRSSVLQVVPFLVSVGADNIFIFVLEYQRLP---RRPGPREVH-----I 736
DB 288 F---IGFRFG-SILCVTPFLVLAIGVDDSYLMAWQRIITCHRRKHARFESVNVLEKNRI 343
QY 737 GRALGRVAPSMLLCSLSSEACFFLGALTMPMPAVRTFALTSGVLVDLLOMSAFVALLS 796
DB 344 TEMFIETGSEITITITITNALAFGIGATTPAAEQIQLFSGNALAVITDFVFTTFFYCALMA 403

Db 175 MLEPARDNGLVGRDYS---STIIISLNREKKQALRVVYNRIKNEIKSNPPEG----- 226
QY 588 OAKLWEEAFLEENRPFQRMAGNFQVFTTAERSLEDEINRTTAEDLPFATSIVIFLYI 647
Db 227 -----VEVVLTDGLGITYKILEMLQNDMRTWA-----VAGIIVILLILLYP 267
QY 648 SLALGYSYSSWSRVMVDSKATLGLGVAVVILGVAAMGPFYSYLGIRSSVLVILQVFPVLV 707
Db 268 -----YKSPFIRMLVPL-----IPLIFGVTM-TLGFMLGLLGLPLD-IATSTVGAMII 311
QY 708 SVGAD-NIFIFVLEYQRLPRRPGEPREVIHGALGRVAPSMLLCSLSEALCFPLGALTFM 766
Db 312 GMGIDVGVHTRRYEE--RNKGRSPEEAERASISGTGKALGAALTTTAGFLAUSLSIL 369
QY 767 PAVR--TFALTSLGVALVDFLQMSAFVALLSLDSKQEARLDVCCCKVQPOLEPPGQG 824
Db 370 PSLKRLSISLVNGLGLAAVNAVIVTPALAIL-----EEFRTKI--LGKGEIISIGE 421
QY 825 EGLL-----LGFQKAYAPFLLHMITRGVILLFLALFGVSLYSWCHISVGLDDELALP 878
Db 422 EGKLGPIFSLGLFIKR-SPWL-----ALLIAMIISGVSLYGATKITTEVRLEKMP 472
QY 879 KDSYLL-----DYFLPLNRYEVCAPVYFTTLYG--NFSSEAGNAICSSA 923
Db 473 MDMEIQALTDIRSEFGQDEITILIKADVDVRDEIVRSILRPEREVKADSYINNPFSE 532
QY 924 GCNNSFTQIOVATEFPQSYLAIPASSWDDFDIDLWLTSPSCCRLLYISGPNKDFCST 983
Db 533 STADV-VIRKYGYIPEDKEKIAKALEGSLV-----SS 564
QY 984 VNSLNCLK---NCMSITMGSVRPSVEQFHKYLPWFLNDRPNIKCPKGGGLAAYSTSVNLT 1040
Db 565 DYSLATVKLGNFMVGTQ-----EDFNKIMEYFEERUKRANFPFGVKA----- 607
QY 1041 DQGVLASRPMWYHKLKNSQDYTEALRAARELAANITADLRKVGTDPAPFVPTITNV 1100
Db 608 -----ELAGE-----SYLNTVLDNL 622
QY 1101 FYEQVLTILPEGLFMLSCLV---PTFAVSCLLGLDLRSLGLNLLSIVMLVDTVGF 1155
Db 623 VNEELGKSTIGTIIIVLVIPAFIRPTVAVAM-----IMPFLGALWJTIY 669
QY 1156 MALMDISYNAVSLNLVS-AVGMSVEFVSHITRSPFAISTKPTWLERAKE-----ATI 1206
Db 670 MGLAGIPFSQ-FLAGVSMIVGLVDYGMHITRP-----LEELKEGKNTPIITAME 720
QY 1207 SMGSAVFAVAMTNLPGIIVLGLAKAQLIQIFFRNLNLTILGLLHGLVFLPVLISYVG 1266
Db 721 SVGPGLIG-ALTAGGFLALITGRLTAIHDFG-----KVLAVG-----IFSSMLAAVL- 768
QY 1267 PDVNPAL 1273
Db 769 ---VTPAI 773

RESULT 37
T33156
hypothetical protein F56H1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33156
R:Graves, T.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid F56H1.
A:Reference number: Z21293
A:Accession: T33156
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-969 <GRA>
A:Cross-references: UNIPROT:O76372; UNIPARC:UPI000017BA34; EMBL:AF067618; PIDN:AAIC19198.
A:Experimental source: strain Bristol N2; clone F56H1
C:Genetics:

A:Gene: CESP:F56H1.1
A:Map position: 1
A:Introns: 49/3; 128/1; 199/3; 236/1; 329/2; 341/2; 361/1; 391/3; 416/1; 465/3; 497/2;
Query Match 2.5%; Score 173; DB 2; Length 969;
Best Local Similarity 18.0%; Pred. No. 0.00079;
Matches 185; Conservative 164; Mismatches 357; Indels 324; Gaps 45;
QY 375 LTTPDVLEWLSAPNSQARSEKAFHQHFGPFRFTNQVILTPNRSYRYDS----- 424
Db 86 IDTDVQPSAHPVSYKSKSWADN-----LLSAINQVAC-YDSPPLMDHLSQ 133
QY 425 LLLGPNPFGIILDLLELLELQRLR-----HLQWSPQARNISLODICYA 473
Db 134 FILEPVNYDAIFNLKFLDKLQMQSNISPLSRDFAFTYRNIS-----VANMEACIS 187
QY 474 P-LNPDNTSLYDCCINSLSLOYFON--NRTLLLLTANQTLMO-----TSQVDW 518
Db 188 PNLRVNCTELDESDIKIVKRTIDNCWKYRTQIFECRDEKCGCCATCRLPENCSSQIMY 247
QY 519 KDHFLYCANAPLTFKDGITALALS CMADYGA PVPFPLAIGGYKGD-----YSEARALI 571
Db 248 D---JFYRLLPKQ-RDTPFLVNTP-----LPFTLTGYITONIPVNVIIYDDLETSI 296
QY 572 MTFSLNNYPAGDPRLAQAQAKLMEAEFLEENRAFQREMGAMFQVTTAERSLEDEINRTTAE 631
Db 297 IDYSKKN-----KPHLKGEFTIIIQNDPFLSKKIVNDNLGP 331
QY 632 DLPI-PATSYIVIFLYISLA-LGSYSSWSRVMVDSK-----ATLGLGGVAV 675
Db 332 NISSQFLAICITIFNFYAKYKSYETRGLLMDVKRDLAALRDSMLALLAAGLVMI 391
QY 676 VLG-----AVMAAMGFYSYLGIRSSVLVILQVVPF-LVLSVGADNIFI 716
Db 392 VVAHSQSLLYAFVVILLALSVGALGVISL--FTDEPPLNLVTFVLLIAGSDDAFL 449
QY 717 FVLEYQRLPRRPGEPREVIHGALGRVAPSMLLCSLSEALCFPLGALTMPAVRTPALTS 776
Db 450 LKSNF-----PKHLNEDTEFTLSHTSFTMTLTCTCFSTIVPFFINITSVIVPRCGLFA 503
QY 777 GLAVILDFELQMSAFVALLSLDSKQEARLDVCCCKVQPOLEPPQGGGLLIGFPQKAY 836
Db 504 GVTVIFNFYFMVVSFLPAFLLIQYRHF-----CFTGKFPYRSVLSHLLY----- 548
QY 837 APFLHMITRG-----VLLFLALFGVSL-YSMCHI-----SVGLDQ----- 873
Db 549 -VLLPVLVQGRYVILSLISVAIGAAITVQGLHLPEYNPLQFTSDMLHFWNNNAER 607
QY 874 -----BLALPKDSYLLDYFLNRYFEVGPVYFVTTLGYNFSSEAGNAICS--SAG 924
Db 608 NPEFVSAXIALPLTSRLV-----NGV-EPIYSLSTFRANATSPLSRDPISLKTAG 657
QY 925 -----CNNPSTQKIQYATEPPEQSYLAIPASSWDDFDIDLWLTSPSCCRLLYISGNKQKF 979
Db 658 DVRKLARFLGTARQLPFPVNHQPK-----FWPERFLWSDKYPFCARGFL----- 700
QY 980 CEPSTVNSL-----NCLKNCMSITMGSVRPSVBEQFHKYLPWFLNDRPNIKCPKGGGLAAY 1032
Db 701 CNMTNPLFSDSYLDFCLRNST-----FLATSYNDTP----- 733
QY 1033 STSVNLTSDGOVLASRRFMAHYHKLKNSQDYTEALRAARELAANITADLRKVGTDPAPREV 1092
Db 734 -----IFDNKTFA--FVGYYTAMLPSTLKYNHRPKQLSK-----SFEMLEMTKPDNGWA 780
QY 1093 PPYITNVFYQYQYTLPEGLFML--SLCLVPTFA-----VSCLLGLDLR 1136
Db 781 PEWMILMSTWFDLLSSIVQDCLSSVGVSLVFAIFAIFIQLKFOQAAVAVTIACVI---TS 837
QY 1137 SGLNLLSIVMLVDTVTCFPMALMDISYNAVSLINLVSAVGSVVEFVSHITRSPFAISTKPT 1196
Db 838 SAIVTLGLGVILGVLEAV-----ILVVLVGLSFDYTLHY--GAALPEHGCG 879
QY 1197 WLERAKEA-----TISMGSAPVAGVAMTNLPGI-----LVGLAKAQLIQIF 1238

Db 880 AEHREATSKGVPVTLAFTSLAGASM--LPALTHAFYQGVFLWISITWTFSTP 937
QY 1239 FFRNLNLT 1248
Db 938 FYLPMLSLT 947
RESULT 38
A:Species: Pyrococcus abyssi (strain Orsay)
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: A75018
R:anonymous, Genoscope
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: A75018
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-746 <KAW>
A:Cross-references: UNIPROT:Q9UY29; UNIPARC:UPI00003456A; GB:AJ248288; GB:AL096836; NII
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1101
C:Superfamily: Archaeoglobus conserved hypothetical protein AF0459
Query Match 2.5%; Score 170; DB 2; Length 746;
Best Local Similarity 19.1%; Pred. No. 0.00091;
Matches 146; Conservative 120; Mismatches 290; Indels 210; Gaps 32;
QY 550 VPPFLAIGYKGDYSEALIMTSLNYPAGDPLRLAQAALWEAFLEEMRAFQRR--- 606
Db 134 MLPPEARNLISGDYS---STIIIVNLN-----REKNQKALVRVYNDIEREIRTFPP 183
QY 607 -----MAGMFQVTFPAAERSLEINRTTAEDLPFPATSYIVIFLYISALGSYSWSRV 660
Db 184 QGVEVLTGDLGYKILELQNDNRWTA---ISGIIVLILYF-----YKSPIRM 233
QY 661 MVDSKATLGGVAVVLGAVMAAMGPFYSILGRSSILVILQVVPFLVLSVGAD-NIFIFVL 719
Db 234 LVPL-----VPLIFGVVW-TLGPMLGLPILD-IATTTVGAMLIIGMIDYGVHVTR 283
QY 720 EYQRLPRRGPPEVHIGALGRVAPSMLLCSISEAICFFLGALTPMPAVRFFALTSLGA 779
Db 284 YTEE--RGKGSLEEAEEAIAETGKALLGNALTITAGFLALSITLPSLKKLSVSLVWG 341
QY 780 VILDFELQSAFVALLSLDSKQEAASRLDVCCCKVQPELPFPQCGEGLLGFFQKAYAPP 839
Db 342 LGLAALNNAVITPALAILSEEFREK-----VMKKEIIAIGGG-----KSKISF 385
QY 840 LLHWITRGV-----VLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRY 893
Db 386 IFSMGLKAIKRSPTALLIALIISGVSLYGASKITTEVRLEKMIPTDL----- 433
QY 894 FEVGAPVYFTTLGYNFSSBAGNAICSSAGCNFPSTQKI-QYATEFFEQSVLAIPASS 952
Db 434 ----PEIQALSDIRSEFGQDEVITLIKADDVRDPTLVRDILRFEREIKADSY----- 482
QY 953 WVDFFIDWLTSPSCRLYISGPNKDFCPSVTNSLNCNCSITMGSVRPSVEQPHKYL 1012
Db 483 -----INNV-----FEQSIADVVIOKYGI 503
QY 1013 PWFILNDRPNIKCPKGGIAAYSTSVNLT---SDGVLA-----SRFMAHYKPLKNSQDYT 1063
Db 504 P---QDKEKISEALEGSSLSVSDYSMTIILKGNFMGVTOSEFNRIMEYFERAIQADFP 560
QY 1064 EALRAABELAANITADLRKPGTDPAPVEFPYITINVFYQYLITLPEGLFMLSCLVPT 1123
Db 561 PGKRV--ELAGE-----SYLNVLNDLVNEELGKISTIGTGLIVVVVFAL 603
QY 1124 FAVSCLLLGLDRLSGLNLLLSIVMLIVDTVGFMALMDISYNAV-----SLINLVS-AV 1175

Db 604 FR-----RPTVSIAMIMPFLG--ALMTIGYMGLAGIPFSQTLAGVWSMIV 647
QY 1176 GMSVEFVSHITRSFAISTKPTWLERAKE-----ATTSMGSAVPAGVAMTNLPGILVL 1227
Db 648 GLGVDYGMHITHRP-----LEELNEGNTPIITALESVGPGIFVG-ALTAGGP--L 696
QY 1228 GLAKAQLIQIFFFRLLNLLITLLGLLGLVFLPVLVSIVYGVDPVNPAL 1273
Db 697 ALLSAQLTAIHDFGRVLAVG-----IFASMPAAVL---VTPAI 731
RESULT 39
B81282
A:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81282
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chiller
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A:Reference number: B81250; MUID:20150912; PMID:10688204
A:Accession: B81282
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-823 <PAR>
A:Cross-references: UNIPROT:Q9PWS7; UNIPARC:UPI00000C1EE7; GB:AL1139078; GB:AL111168; NII
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1373
Query Match 2.2%; Score 154; DB 2; Length 823;
Best Local Similarity 18.3%; Pred. No. 0.014;
Matches 168; Conservative 153; Mismatches 351; Indels 244; Gaps 34;
QY 393 EKAFHDQHGFPPFTNQVILTAPNRSSRYRYSLLGPKNPSGIILDL---LELLEQER 450
Db 76 EKPFSENELAKLKKLHELEKAP-----LVERVFS-IINAPLQSSQNTDLKEL 123
QY 451 LRHL-QVWSPQAQRNISLQDICYAPLNPDTSLDYDCCINSLLQVFNQNRTELLLTANQTL 509
Db 124 LKNIPNIESQDINRTKAQNEILNPSFFYKNNIISKDGKVTGLIILYKPTVY-----NNLI 178
QY 510 MGOTSOVDMKDHFLYCANAPLTFKDGITALALSMDYAGAPVFPFLAIGGYKGYSEABA 569
Db 179 EKRDLAYTDEK-----KQIRLAI-----KEHQDKQK 205
QY 570 LIMTFSLN-----NYPAGDPLRLAQAALWEAFLEENRAFQRRMAGMFQVTFPFAERSLE 622
Db 206 VITKSLDTIKSIVRNYEKNKDTL-----YLGGVSMIA----- 238
QY 623 DEINRTTAEDLPFPATSYIVIFLYISLAL-GSYSSWSRVMSVDSKATLGLGGVAVVLGAVM 681
Db 239 DDMIAIYKSDLVLYGVGS--LVFL-LGLALYFFFSWRVFLP-----LFICFISLS 286
QY 682 AAMGFYSYLGRSSIV---ILQVDFVLVSGADNIFIVLEYQRLPRRPGSPREHVIGR 738
Db 287 AASGVFALLNQITVISSNYVALVLIITLSVVVHLIITHFIESTQNHKPAKVE--RIVLET 344
QY 739 ALGRVAPMLLCSLSEACFFLGALTPM-PAVRTFALTS---GLAVILDPLLOMSAFVAL 794
Db 345 LLAKANPS-LVAIVTMTIGFFSLILSNIEPIIKLGIMMSIGLGLALIFYFLFLASILVLL 403
QY 795 LSLDSKQEAASRLDVCCCKVQPELPFPQCGEGLLGFFQKAYAPFLHWTGRVLLAFL 854
Db 404 KPNVHKKEKFNLLAFCAK-----TALDPKRRMIYISVLAIL 444
QY 855 ALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPVYFTTLGYNFSSBA 914
Db 445 ALIGISKLRV-----ENSFV-----NYFKDGSBI-----KK 470
QY 915 GMAICSSAGCNFPSTQKIQYATEFP-----BQSYLAIPASSWVD-- 955


```
Db 471 GLLVI-----DKNLGGLTFLVLIIRPNKNNDQNASDTLSDSESEFENLATQTYWFDK 525
QY 956 -----DFIDWLTPSCCRLYISGPNKDFPCPTVNSLNCNCKMCMSTIGWSVRPSVE 1006
Db 526 KTRIAKKVHEFL-----NKEFVGSVLISLSL-----LTGKNINDK 563
QY 1007 QFHXYLPWFLNDRPNIKCPKGGIAAYSTVNLTSQGVILASRFMAYHKPLKNSQDYTEAL 1066
Db 564 ELDDPALAFLENLPAKFKQDLSPF---VSIENNELRFMSRIVSDPNLRRNEFLDLK 620
QY 1067 RAARELAANITADLRKVGTDPAFEVFPPTITNVFYEQYLITLPEGLFMLSCLVPTPAV 1126
Db 621 KQLNELVRKDGVEVQ-----ITGMVLNNMLQSLFSSQFDTLIPVILAIFI 667
QY 1127 SCLLGLDLRSGLLNLISWMLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHIT 1186
Db 668 LPIIVPRDLKFSVAAILVNVIPLSVFPALMGLLIGIPDMMSITIAAIGVDDDAIHY 727
QY 1187 RSFAISTKPTWLER-KEATISMGSAVFAGVAMTNLPGILVLG---LAKAQIQLIFFPRL 1242
Db 728 YRFREIKNKSLEEAIMSHLSGALY---YTTIS--IVLGSVMMSSNPFTIYFGI 781
QY 1243 NLLITLLGLLHVL 1258
Db 782 LTVFVMIILLSGSLF 797

RESULT 40
T07712
probable ABC-type transport protein T23J7.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C:Accession: T07712
R:Queix, P.; Choine, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigou
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15793
A:Accession: T07712
A:Molecule type: DNA
A:Residues: 1-1011 <QUR>
A:Cross-references: UNIPROT:Q9STU0; UNIPARC:UPI00000A52C0; EMBL:AL049746; GSPDB:GN00061;
A:Experimental source: cultivar Columbia; BAC clone T23J7
C:Genetics:
A:Gene: ATSP:T23J7.60
A:Map position: 3
A:Introns: 123/3; 201/1; 257/3; 319/3; 351/3; 419/3; 473/1; 548/3; 588/1; 613/3; 643/3;
C:Keywords: ATP; p-loop

Query Match 2.2%; Score 149; DB 2; Length 1011;
Best Local Similarity 20.7%; Pred. No. 0.04;
Matches 114; Conservative 83; Mismatches 237; Indels 116; Gaps 25;
QY 768 AVRTFALTGLAV-ILDFFLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPPGQGR 826
Db 223 ALDTTGPTFLAVMGFVLQISLITEREL--KLQAM----- 259
QY 827 LLLGRFQKAYAPLLHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSVLLDY 886
Db 260 TWMGVEDTAY---WLSWLTWEGILTALISALLVLFQGMFQDFFL-----KNSFPVVF 309
QY 887 FLFLNRYFVGAPVYFV-----TTLGNFSSAGMNAICSSAGCNFSFTQKIQY 936
Db 310 LLFLMLQFNQLIGLAFMLSAFISKSTATTVGPPVFLVGVFTQLATSSG---PPYACKYR 366
QY 937 ATE-----PPEQSY-----LAIPASSWDDFDIMWLTSPSSCCRLYISGNKQKFCPSTVN 985
Db 367 MIRALMSLFPNFTSQGLKLADATSTPDQPGISMSKRAEC-----GPNDDTGCVLITN 420
QY 986 SL-----NCLKNCMSITWGSVRPSVEQHKYLPWFLNDRPNIKCPK 1026
Db 421 DIYLMGLTGTFWFLVLYFDNITPNA-----SGVRKSIYFLK--PGYWTGKGNRVEE 473
QY 1027 GGLAAYSTSV---NLTSQDQ-VLASRFMAYHKPLKNSQDYTEALRAARELAANITADLR 1081
```

```
Db 474 GGICSGISGVPVVDHITPDDEDVLEETLVKQHSMEGLVDPNVAVQ-IRGLAKTYPGTTK 532
QY 1082 ----KVPGTDPAFEVFPPTITNVFYEQYLITL-PEGLFMLS-CLVPTPAVS---CLLL 1131
Db 533 FCCCKCKKTSP-FHAKGLWMMIAKQDLFCGLPGNAGKTTTINCITGLTGPVTVGGDALLY 591
QY 1132 GLDLRS--GLLNLGSLWMLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSF 1189
Db 592 GNSIRSSVGSNIRKMWIGVCPQ---FDILMDALSGBEHLKLFASIKGLPPSSINSVVEKS 648
QY 1190 AISTPTPTWLERAKEATISMGSAVFAGVAMTNL--PGILVLG--LAKAQIQLIFFPRLNLL 1245
Db 649 LAEVKLTEAGKIRAGSYSGMKRRLLSAVSLIGDKPLVFLDPVCNIIYSDVYVLFKSLFL 708
QY 1246 ITLLGLLHGL 1255
Db 709 FSCVLVIRGL 718

RESULT 41
AG3455
multidrug resistance protein mexB [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AG3455
R:DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3455
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1051 <KUR>
A:Cross-references: UNIPROT:Q8YF93; UNIPARC:UPI0000058143; GB:AE008917; PIDN:AAL52810.1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11629
A:Map position: 1
C:Superfamily: acriflavin resistance protein

Query Match 2.2%; Score 149; DB 2; Length 1051;
Best Local Similarity 19.8%; Pred. No. 0.042;
Matches 221; Conservative 162; Mismatches 417; Indels 318; Gaps 56;
QY 290 ILCSPVAVTIL--LVGFRVAPARKSKMVDPKGTSLSKLSFSTHTLLGQFFQGWGTW 347
Db 10 VPAAVISIVLVLAGLICIRILFVAQYPELTTPQVVSA-----TYPGASAEVAQT 60
QY 348 VASWPLTILVLSVIPV-----ALAGLVFTLT---TDPVELMSAPNS---QARSEKA 395
Db 61 VAA-FLEQOINGVENLWQSSLSGSGTQMLTFTVTFALGTDPPQATINVNNRVQRTATSSLP 119
QY 396 FHDQHFQFPF--RTNQVI--LTAPNRSSVRYDSLLGP-----KNFSGILDLDLL 442
Db 120 QEVQRGLVTVDRKSTILGMVAMFATDRYDRTYGVNALLNVDDKELPGVGVQVLLG 179
QY 443 ELLEQLERLHQLQVMSPEAQRNISLQDICYAPINDPNTSLYDCCINSLLQYQNNRTLL 502
Db 180 NI-----DYSNRVW-----LRPDLAQYNLTFSVSAAIQEQNA--- 213
QY 503 LTANOTLMGQTSQVDMKO---HELYCANAPLTFKDGTA---LALSCMADYGAVPFPFLA 555
Db 214 ----QPAAGRFG--DQDPHAGFPFTTATQGRLPDAAAFENIILRSSQNAATL----- 262
QY 556 IGGYKGYDYSEABALIMTFSLNNYPAGDPRLAQAALWBEA--FLBEMRAFQRRM----- 607
Db 263 ----RLKQVARVELGTESVYVNSNLNGTPAVPIAIVLPQGANALNTMELIQNRNELKAS 318
QY 608 --AGM-FQVTFPAERLEDEINRTTAEDLPFIPATSYVIFLYISALGSSYNSWRVWDS 664
Db 319 FPAIDYAIPTFTTKVKSIEBHVHTFIEAILVLVVFIFL-----QNW----- 364
```

QY 665 KATILGGVAVVLGAVMAWGPFSSYLGIRSSIVLLOVPFLVLSVG--ADNIFVFLEYO 722
:
:
:
Db 365 RPLR-----IPVIAPDISIIIGTAGGVLYGFSGINLTLLFGLAIGLVDDVAIVLENVE 419
:
:
:
QY 723 R-LRPRLPGPREVHIGRALGRV-----APSMLLCSLSEAI CFFLGALTPMPAVRTFA LTS 776
:
:
:
Db 420 RIMTTEKLSPKAAI-KANGEVTVPIAIVLVCAVFPVAFPMGGLVGEM--YKQFAVTI 476
:
:
:
QY 777 GLAVIDLFLLQWSAFVALLSLKSKQEASRLDVCCCPQLPPEPPQGEGILL-IGFPOKA 835
:
:
:
Db 477 AISVTLSELGVALTLPALCALIK-----PCGHHEPILPRFIENRA 516
:
:
:
QY 836 YARFELLHWTRGV-----VLLLFLALFGVSLSYMCHISVGL-----DO----- 873
:
:
:
Db 517 FER-VTSGYTRGVRFELKRATITGLIIFAGLLGSTYYLPERVPSGLLPDEDQGLFEGVA VL 575
:
:
:
QY 874 --ELALPKOSYLDY-----FLLNRHYFEVGAPVVV----- 903
:
:
:
Db 576 PPAASLERITVVLQDSVENIRKNPAVDNVFAVSGFDLLSGLKTSGATWFMILDKWKERT 635
:
:
:
QY 904 -----TTLGNFSESAGMNAICSSAGCNFSFT-----QKIQA 937
:
:
:
Db 636 TPDADARNLRPTMGNNAGIKOGMWLAFFFFPPIMGSLSTTGGEFYVQDRTGGVESLTOA 695
:
:
:
QY 938 TEFPESGLYAIPASSWDVDFIDWLTPSSCCRILIYSNPNDKFCSPTSVNSLNCLNCMSIT 997
:
:
:
Db 696 TKLITEANAARPELOGVRTTFDPVPOYDIQ-----DREKAKMGV-PINSVETAMQAT 749
:
:
:
QY 998 MGSVRPSVEQFHXY-----LPWFNLDRPNIK-----CPKG-----LAAYSTSVMILT 1039
:
:
:
Db 750 FGS L-----VYNDFTLYGRNYQNVLQSEAERFRDPCDLKHVFRADSGSMIFDALVTVKRI V 807
:
:
:
QY 1040 SDGOVLASFWAYH--KPLKN-SODYT--EALRAARELANAITADLRKVGTDPAPE--- 1091
:
:
:
Db 808 GPDQL--ERENAFNAAKVTGNPAGYTSGDAIKAMOEVAQAQLPQGYQTAWTGSAYQEV S 865
:
:
:
QY 1092 -----VFPTYTNVRYEQYLI TLPEGLFMLSLCLVPTFAVSCLLGLDLRSRG 1138
:
:
:
Db 866 TSGTGSQMIFGLIMVFLILA AQYRW--SLP-----LAVITAVPPAIFCALLATDLR-G 917
:
:
:
QY 1139 LLNLLSTVMILVDTVGPMAWDISTYNNAVSLINLAVSGMSVEPVSHITSFAISTKPTWL 1198
:
:
:
Db 918 LT N-----DVYFQ-----IGLVTLIGLAVKNAILIV-EFAV-----L 948
:
:
:
QY 1199 ER-----AKEATISMGS AVFAGVANTNLP Gil-VLGLA 1230
:
:
:
Db 949 ERESGKAIEAAAARAIRFPPIVMTSLTILGWVPLA 986
:
:
:
RESULT 42
H84301
hypothetical protein Vngl1476c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84301
R:Ng, W.V.; Kennedy, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable r; Leithauer, B.; Kellner, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laskey, S.; Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A:
Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; PMID:20504483; PMID:11016950
A:Accession: H84301
A>Status: preliminary
A:Molecule type: DNA
A:Residue: 1-746 <STO>
A:Cross-references: UNIPROT: Q9HPT7; UNIPARC:UPI00000638D0; GB:AEO04437; NID:g10580974; E C:Genetics:
A:Gene: VNG1476C

Query Match 2.1%; Score 148.5; DB 2; Length 746;
Best Local Similarity 21.6%; Pred. No. 0.029;

[illegible]

RESULT 43

AC0423

multidrug efflux protein [imported] - *Yersinia pestis* (strain C092)

C/Species: *Yersinia pestis*

C;Date: 02-Nov-2001
C;Accession: AC042323

C; ACCESSION: AC0423
P: PARKHILL, J. : Wre

K; FALKHILL, C.; WOOD
deno-Tarraga. A.M.:

il, M.: Rutherford,

Nature 413, 523-527

A;Title: Genome seq

A;Reference number:

A;Accession: AC0423

A;status: preliminary

Qy 1271 -----PALAEQKR 1279
Db 702 EGHQLTKEIQEIDHEQKK 720

RESULT 45
AF0561
A:Title: acriflavin resistance protein B [imported] - Salmonella enterica subsp. enterica serovar
A:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AF0561
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AF0561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <PAR>
A:Cross-references: UNIPARC:UPI000005A1FB; GB:ALU513382; PIDN:CAD04960.1; PID:G16501747;
C:Genetics:
A:Gene: STY0519
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 141.5; DB 2; Length 1049;
Best Local Similarity 18.1%; Pred. No. 0.14;
Matches 145; Conservative 110; Mismatches 243; Indels 303; Gaps 32;

Qy 624 EINRTAEDLPFATSYIVIFLYISLALGSYSSWSRVVDSKATL--GLGGVAVVLG--A 679
Db 339 EVVKLVE-----AIIIVFLVWVLFQNF-----RATLIPTIAVPVVLGTF 381

Qy 680 VMAAMGFSSYLGRSLVILQVVPFLVSVG--ADNIFIVLEYQRLPRRPGEPREHVIG 737
Db 382 VLAAGF-----SINTLTMF--MVLAIGLLVDDAIVVVENVRVMTTEGLPPKEATR 432

Qy 738 RALGR-----VAPSMCLSLSEAI CFFLGALT PMPAVRTFALTSLGLAVILDFLLQMSAFV 792
Db 433 KSMGQIQGALVGAMVLSAVFTPMAFFGGSTGAI--YRQFSITIVSMAISVLVALIITP 490

Qy 793 ALLSLDSKRQESRLDVCCCKPQELPPPGQEGLLGFFQKAYAPFLLHWI----- 844
Db 491 ALCA-----TWLKPVAKGDHGGKGGFGWFLNLFDKSTHHTYDTSVGNILR 536

Qy 845 TRGVLLLLFLALP-----GVSLSYMSCHISVGLDQELALPKDSYLL 884
Db 537 STGRYLLLLIIVVGMAIYFLRPLSSFLPDEDQGVFL--TMVQLPAGATQERTQKVLDEVT 595

Qy 885 DYEL-----FLNR----- 892
Db 596 DYYLNKEKANVESVFAVNGFGFAGRGONTGIAFVSLKOWADRPGEKVKYALTQRTAATP 655

Qy 893 -----YFEVGAPVYFVTTLTGYNFSSEAGNNAICSSAGCNFFSTQ----- 932
Db 656 SQIKDAMVFAFNLPALVELGTATGDFE-----LIDQAGLGHKLTQARNQLGEVAK 708

Qy 933 -----KIQATEPEQSYLAIPAS-----SWVDDFDLWLTTP 963
Db 709 YPDLGVVRPNGLDTPQFKIDIDQEKAAQALGVSISDINTTLGAANGSGSYVNDFFID--R 765

Qy 964 SSCRLYISGNPKDKFCPSVTNSLNCNLCNMSITMGSVRPSVEQPHKILPFWFLNDRPNIK 1023
Db 766 GRVKVYVMEAKYRMLPDDIND-----WYVSGSDGQM 798

Qy 1024 CPKGGAAAYST-----SVNLTSDGQVLSRFRMAYHKPLKNSQDYTEALRAARELANIT 1077
Db 799 VP---FSAFSSRWYSGPRLERYNGLPSMELIQQAAAPGKST---GEAMAMBEELAKLP 852

Qy 1078 ADLRKVPGTDPAFEVPPYTTITNVFBOYILT-----ILPEGLFMLSCL----- 1120
Db 853 SGI-----GYDWTGMSYQERLSGNQAPALYALISLIVVFLCLAALVESWSIPP 899

Qy 1121 -----YPTFAVSCLLGLDLRGLMLLSIVMLIVDTVGFMAIWDISYNVSLIN----- 1170
Db 900 SVMLVVPLGVIGALLAA--TFRGLTNDVYFQVGLLTIGLSA-----KNAILIVEFAKOL 952

Qy 1171 -----LVSAYGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMT 1219
Db 953 MDKEGKGLVEAMLEAVRMRLRPILMTSLAPMLGVP-----LVISSG-----AGSGAQ 1000

Qy 1220 NLPQILVL--GLAKAQILQIPIFF 1239
Db 1001 NAVGTGVLGGWVTATVLAIFF 1021

RESULT 46
D85757
A:Title: efflux pump Z2508 [imported] - Escherichia coli (strain O157:H7, substrain BDJ9
A:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85757
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1027 <STO>
A:Cross-references: UNIPROT:Q8X7E2, UNIPARC:UPI000000D0A77; GB:AE005174; NID:G12515493;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z2508
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 140.5; DB 2; Length 1027;
Best Local Similarity 18.3%; Pred. No. 0.16;
Matches 149; Conservative 125; Mismatches 261; Indels 281; Gaps 38;

Qy 576 LNNTYA-----GDPRLAQAKLWEEAFLEENRAFQRMAGFMQVTFTHRSLEDBIN 626
Db 280 LNHPSAGVAVMLSPGANALNTATLVKDIAE-----FORNMPQGYDIAYPKDST---BFI 332

Qy 627 RTTAEPLPFIATSYIVIFLYISLALGSYSSWSRVVDSKATLGLGGVAVVLGAVMAAMGF 686
Db 333 KISVED--VIQTLFEAIVLVVCMYLFLOQLRATLIIPALA-----VPVLLGTGVLAL 384

Qy 687 PSYLGIRSSLVILQVVPFLVSVG--ADNIFIVLEYQRLPRRPGEPREHVIGRALGRVA 744
Db 385 FGY-----SINTLTIFA--MVLAIGLLVDDAIVVVENVERINRDKGLPAREATERSMGEIS 438

Qy 745 PSMLICSLSEAI C-----FLGALT PMPAVRTFALTSLGLAVILDFLLQMSAFVLLSDSK 800
Db 439 GALVAIALVLSAVFLPMAFFGSGTGV--YRQFSITIIISAMLSVVVALTLTPAL----- 491

Qy 801 ROEASRLDVCCCKPQELPPPGQEGLLGFFQKAYAPFLLHWITRGVVLVLLFLALFGVS 860
Db 492 -----CGSVLQHVPPHKKG---PFGAPDFRFRRTEDKY--QRGYIVYLRRAARTMG 537

Qy 861 LYSNMCHISVGL-----DQ-----ELALPK-----DSYLLDYFLFLNR 892
Db 538 LYLVLGGGHALMMWKLPGSFLPTEQGEIMVQYTLTPAGNTAARTAEVNRQIVDWFL--INE 596

Qy 893 YFEVGAPVYFVTTLTGYNFSSEAGMN----- 917
Db 597 --KANTDVI F--IVDGFSG--SGQNTGMAFVSLKNWSQRKGAENTAQAIALRATKELGTI 652

Qy 918 -----AICSSAGCNPF-----TKIQYATEPPE----- 942
Db 653 RDATVFAMTPPAVDGLGQSGNGFTFELLANGGTDRETLQMRNQLIEKANQSPELHSVRAN 712

QY 943 -----QSYLAIP-----ASSW-----VDDFDMLTSSCCRLYISGP 974
Db 713 DLPQMPQLQVDIDSNKAVSLGLSLNDVDTLSSAGGTYYNDFID---RGRVKVVIQGD 769
QY 975 NKDKFCPSTVNSLNCNKMSITMGSRVPSVEQFKYLPWFLNDRNPKCPKGLAAYST 1034
Db 770 SEFRSAPSDLGK-----WVFRGSDNAMTP-----FSAFAT 799
QY 1035 SVNLTSDDGVL A-SRFMAYHKPLKNSQDYT--BALRAARELANITA-----DLRK 1082
Db 800 TRWLYGPERLVRVNGSAAYEIOGENATGSSGDAMTKMELANSLPAGTTWAWSGLSLOE 859
QY 1083 VPCTDPAFEVPPYITNVF-----YEQY-----LTILPEGLFMLSCLVPTFAVSCLL 1130
Db 860 KLASGQALSILYAVSILVWFLCLAALYESWSPFVILVPLG-----L 902
QY 1131 LGLDLRSGLNLLSIV---MILVDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHITR 1187
Db 903 LGAALAAWRDLNDVYFQVALLTIGLS-----SKNALIVFEPAEA-----VAE 948
QY 1188 SFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPG 1223
Db 949 GYSLRAALRAAQTRLRPIIMTSLAP--IAGGNAAG 982

RESULT 47
H90861
probable multidrug-efflux transport protein protein ECe1864 [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90861
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90861
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1027 <HAY>
A:Cross-references: UNIPROT:Q8X782; UNIPARC:UPI00000D0A77; GB:BA000007; PIDN:BA035287.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECe1864
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 140.5; DB 2; Length 1027;
Best Local Similarity 18.3%; Pred. No. 0.16;
Matches 149; Conservative 125; Mismatches 261; Indels 281; Gaps 38;
QY 576 LNNYPA-----GDPLAQAQKWEAPLEMRARFORRMAGMFOVTFATERSLEDEIN 626
Db 280 LNHPSAGVAVMLSPCANALNTATLVKDKIAE-----FQRMPOGYDIAPKDSI---EPI 332
QY 627 RTTAEDLPATSYIVIFLYISLALSYSSWVRVMSKATLGLGGVAVVLGAVMAAMGF 686
Db 333 KISVED--VIQTLFEALVLVVCVMYFLQNLRAATLIPALA-----VPVLLGTGFGVAL 384
QY 687 PSYLGIRSLVILQVFFLVSVG--ADNIFIVLEYQLPRPGRPREVHIGRALGRVA 744
Db 385 FG-----SINTLTFLA-MVLAIGLAVDDAIVVENVERIMRDKGLPAREATEKSMGETS 438
QY 745 PSMLLCSLEAICF-----FLGALTTPMPAVRTALTSLGAVILDFLLOMSAFALLSLDSK 800
Db 439 GALLVALVLSAVFLPMAPFGSGTG-YYRQSITIIISAMLSVVVVALTTTAL-----491
QY 801 ROEASRLDVCCCKVQELPPPGGEGLLGFFQKAYAPFLHLWITRGVLLFLALFGVS 860
Db 492 -----CGSVLQHVPHKKG-----PFGAFDRFYRTEDKY-QRGVIVLRLRAARTWG 537
QY 861 LYSMCHISVGL-----DQ-----ELALPK-----DSYLLDYFLFLNR 892

Db 538 LYLVLGGGMALMMWKLPGSLFPTDQGEIMVQYTLTPAGATAARTAVNRQIVDMFL-INE 596
QY 893 YFEVGAPVYFVITLTYGYNFSSEAGMN----- 917
Db 597 --KANTDVLIP-TVDFGSFSG-SGQNTGMAPVSLKNWSQRGAENTAQATALRATKELGTI 652
QY 918 -----AICSSAGCNSFP-----TOKIQYATEFPE----- 942
Db 653 RDATVPAMTPFAVDGLGQSGNFTFELLANGGTDRETLLQWRNQLIEKANSQSPHLSVRAN 712
QY 943 -----QSYLAIP-----ASSW-----VDDFDMLTSSCCRLYISGP 974
Db 713 DLPQMPQLQVDIDSNKAVSLGLSLNDVDTLSSAGGTYYNDFID---RGRVKVVIQGD 769
QY 975 NKDKFCPSTVNSLNCNKMSITMGSRVPSVEQFKYLPWFLNDRNPKCPKGLAAYST 1034
Db 770 SEFRSAPSDLGK-----WVFRGSDNAMTP-----FSAFAT 799
QY 1035 SVNLTSDDGVL A-SRFMAYHKPLKNSQDYT--BALRAARELANITA-----DLRK 1082
Db 800 TRWLYGPERLVRVNGSAAYEIOGENATGSSGDAMTKMELANSLPAGTTWAWSGLSLOE 859
QY 1083 VPCTDPAFEVPPYITNVF-----YEQY-----LTILPEGLFMLSCLVPTFAVSCLL 1130
Db 860 KLASGQALSILYAVSILVWFLCLAALYESWSPFVILVPLG-----L 902
QY 1131 LGLDLRSGLNLLSIV---MILVDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHITR 1187
Db 903 LGAALAAWRDLNDVYFQVALLTIGLS-----SKNALIVFEPAEA-----VAE 948
QY 1188 SFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPG 1223
Db 949 GYSLRAALRAAQTRLRPIIMTSLAP--IAGGNAAG 982

RESULT 48
S39630
multidrug-efflux transport protein B - Pseudomonas aeruginosa
N:Alternate names: multidrug resistance protein B
C:Species: Pseudomonas aeruginosa
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S39630
R:Poole, K.; Heinrichs, D.E.; Neshat, S.
Mol. Microbiol. 10, 529-544, 1993
A:Title: Cloning and sequence analysis of an EnvCD homologue in Pseudomonas aeruginosa:
A:Reference number: S39629; MUID:95058196; PMID:7968531
A:Accession: S39630
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1046 <POO>
A:Cross-references: UNIPROT:P52002; UNIPARC:UPI000016FC9A; GB:L11616; NID:G438852; PIDN:
A:Experimental source: strain CD10
C:Genetics:
A:Gene: mexB
C:Function:
A:Description: probably involved in secretion of the siderophore pyoverdine
A:Note: inducible under conditions of iron limitation
C:Superfamily: acriflavin resistance protein
C:Keywords: transmembrane protein; transport protein
F14-30/Domain: transmembrane #status predicted <TM1>
F343-359/Domain: transmembrane #status predicted <TM2>
F370-386/Domain: transmembrane #status predicted <TM3>
F397-413/Domain: transmembrane #status predicted <TM4>
F442-458/Domain: transmembrane #status predicted <TM5>
F472-488/Domain: transmembrane #status predicted <TM6>
F541-557/Domain: transmembrane #status predicted <TM7>
F874-890/Domain: transmembrane #status predicted <TM8>
F898-914/Domain: transmembrane #status predicted <TM9>
F973-989/Domain: transmembrane #status predicted <TM10>
F1014-1030/Domain: transmembrane #status predicted <TM11>

Query Match 2.0%; Score 140.5; DB 2; Length 1046;
Best Local Similarity 20.5%; Pred. No. 0.17;


```

Db      910 GALLATSMR-GUSNDVFFQVGLLTIGLSA-----KNAILVFEKELHQKGVIEAAI 963
QY      1173 SAVGMSVEFVSHITSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVL-GLAK 1231
Db      964 EACRMELRPIVMTSLAFILGVVP-----LAISTG-----AGSGSHAIGTGVIGEMVT 1011
QY      1232 AOLIQIPFRL 1242
Db      1012 ATVLAIFWVPL 1022

RESULT 50
Db3206
probable RND efflux transporter PA3522 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83206
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83206
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1053 <STO>
A:Cross-references: UNIPROT:Q9HY87; UNIPARC:UPI00000C5A51; GB:AE004772; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3522

Query Match      2.0%; Score 139; DB 2; Length 1053;
Best Local Similarity 20.4%; Pred. No. 0.21;
Matches 235; Conservative 141; Mismatches 394; Indels 384; Gaps 60;

QY      340 PFQ-GWGTWVASWPLTILVLSVIP-----VVALAAGLVFTELTTPVELWSAPNSQARSE 393
Db      29 PFQPLSEYPVPTPTVQTAVPGANPDVIAQTAVTPLEQAITG-VEGMLYSSQSATD 87
QY      394 KAF-----HDQHFQFF-----RTNQVILTAPNR-----S 418
Db      88 GRMILITTFDHDIDPDMAQIQVNRVSLSPDEVRQGVVQTQKTSPIILMVHLLSP 147
QY      419 SVRYDLSLLGPKNFSGILDLLELLEQLERL-----RHLOVW-SPE--AQRNLSQD 469
Db      148 EORYDPLYI--SNYALQVRDELLRLPGISDVVWVGAGEYSMKRLMDPLDPLIAARGITAGE 205
QY      470 ICYAPLN-----PNTSLYDCINSL--LQVFQNNRTLLILITANQTLMGQTS 514
Db      206 VTAAREQNVQVAGAVGQAPDSTAFAQVTVNTLGRLSDEEQFGDIIVRTGAD---GQVT 262
QY      515 QVWDKDHLYCANAPLTFKDGITALALSCMADYGAVPFPPLAI--GGYKGYVSEBALIM 572
Db      263 RL-----RDVARIENGGDAYALRSLLD-GEPAVALQIIQSPCANALDTAEAVRATV 312
QY      573 TFSLNYPAGDPRLAQAKLWEAFLEEMRAFQRMAGMFQVTFTAERSLEDDINNTAED 632
Db      313 ARLEGNFPAG-----LSARIA--YDPTVFVRASLQTV--ATTLL 348
QY      633 LPFATSYIVIFLYISALGSSRSRVNVDKAT-LGLGGVAVVLGAVMAAGFPFSYLG 691
Db      349 AILLVVVVVFL-----RSWRASLIPLMAVPVSLVGTFF---AVMHLMGF----- 390
QY      692 IRSSVLIVQVPLVLSVG--ADNIFIVLEVQRLPRRGPBPREVHIGALGRV-APSM 748
Db      391 ----SLNTLSLFG-LVLSIGIVDDAIVVENVERHIEN-GEPPLOAARRAMDEVGPIMA 445
QY      749 LCSLSEALCF----FLGALTPMAVTRTEALTSGLAVILDPLQMSAFVALLSLDSKQEQAS 805
Db      446 ITSVLAAVEPIPTAFUSGLQG--BFYQFALTIAISTLSALNSLTLSPALAGL----- 496

```

```

QY      806 RLDDVCCVKPQBLPPPGQGEGLLLGFFQKAYAPFLHMITR----- 846
Db      497 -----LLRPR---DAG---GAVAGRFOR-----LLQVLGRPLRNAPEAYGNVAKVRV 539
QY      847 -GVVLLFLALRGVSLYSCHISVGLDQELALPKDSYLLDDYFLFLNRYEYVG----- 897
Db      540 SGLALVVGGLLGLTW-----VGFP---AVPP-----GFVPMQDKYLVGLVIAQLPNGA 584
QY      898 -----APVYFVTTLGYNPFSSBAGNNAICSSAGCNFSTQKIQTATEPPEQSYLAIPAS 951
Db      585 SLERTDAVVRQMSRIGLD---EPGVESVVAAPGLSVNGFVNVVNAAMF-----FMLDPPE 637
QY      952 SWVDDFIDWLTPSSCCRLYISGNKDKFC-----PSTVNSLNCNKCMSTMG 1000
Db      638 S-----RTSADLGAALAGRLQARPASIPDGFLGVFPVPPVPGLG-----TIGG 681
QY      1001 VRPSVEQFHKYLFWFLNDRPN-----IKCPKGG-LAAYSTSVNLT----- 1039
Db      682 FPMQVE-----DRGGAGLEALARQTVLMMKATESGQLGLMTSPDINAPQLEV 730
QY      1040 -----SDGOVLASRFWAY-----HK----- 1054
Db      731 VDTRFKVKSQVRLADVFEALQVLSLYINDPNRPGRTYKVTQAQADAPHRQAEIGRL 790
QY      1055 -----PLKNSQDYTEALRAARELAAN--ITADLR--KVPGTDP-----AP 1090
Db      791 QVRNAAGAMLPLSSFTVTPSSGPDRIHYNGYPSADISGGALPGVSSQAVAMLERLAG 850
QY      1091 EYVP-----YTTNVFEQVLTILPGLFMLSILCVPTFAVSCLLIGLDRSGLN---LL 1143
Db      851 EVLPEGMFTFETDLTYQOKLA-GNSALFIFPLCVL-----LAYLILAAQYNSWLLPLAVLL 905
QY      1144 STVMILVDTVGFMALWDISYNAVSILNLVSAGMS-----VEFVSHITRSFAISTKPT 1196
Db      906 IVPMLCLSAIAGVMLVGGDNVVFVQIGLVVLGLAKNAILLIVEPARTLEAGA----- 959
QY      1197 MLERAKEATISMGSAVFAGVAMTNL---PGILVLGLAKA-----QLIQIPFRLMLLT 1247
Db      960 ---RALEAVVEACRLRLRPILMTSLAFTAGVVPLVWASGAGAEQMGVAVFAGMLGVT 1016
QY      1248 LLGLLHGLVFLPVI 1261
Db      1017 LFGLELTPVFFVLV 1030

RESULT 51
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An-
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1057 <KUR>
A:Cross-references: UNIPROT:Q8YSE5; UNIPARC:UPI00000CE6AC; GB:BA000019; PIDN:BA074842.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all3143
C:Superfamily: acriflavin resistance protein

Query Match      2.0%; Score 138; DB 2; Length 1057;
Best Local Similarity 18.8%; Pred. No. 0.25;
Matches 197; Conservative 141; Mismatches 323; Indels 388; Gaps 51;

QY      406 RTNQVILTAPNRSYR-YDSLGLGPKNFSGLD-----LDLLELLELLEQLERHLOW-SP 459

```



```
Db 133 QSNWILLAMGLYSQQEQYNTVFLSNYADQYIVDALKRLDGVGEARIIGERRARYAMRLMDP 192
Qy 460 E--AQRNLSQDICVAPLNPN-----TSLYDCCINSLLQYFQNNRTLLLTAN 506
Db 193 NRLASRKLTAQVIDA-INEQNIQVGVQIQGPSLDP-----QWYQ-----IDLQAR 239
Qy 507 QTLMGQTSQVDWKHFLYCANAPLTFKDGDTALALSCM--ADYGAPVF-PFLAIGYGGKD 563
Db 240 GRLSEATEFADM-----VIKTSQDGTILKLDVGRAGELGAENYSSFLRFRNGE-- 288
Qy 564 YSEAEALWTSLNNYPAGDPRLAQAKLWEEAFLEEMRAFQRMAGM-----FQVT 614
Db 289 -----VGIGFPTPGNA-----LEVKAVTEMARLAQSPPGKQYQA 328
Qy 615 FTAERSLED---EINRTTAEDLPIPATSYIVIFLYISALGSYSWSRWVDSKATLGLG 671
Db 329 FDTTLFVEESLAELVVKTLIEALILV--LVIFEL-----QDWRITLIP----- 370
Qy 672 GVAVVLGAVMAMGFFSYLGIRSSVLIVQVPPFLVLSVG--ADNIFIFVLEYQRLPRPG 729
Db 371 -----VITITPLTIGTFAFIKVFGFINTLTTLFGLTLATGLVVDAILIIVENISRLIEQG 426
Qy 730 -EPREV---HIGRALGRV-APSMLCLSLSEALCFELGALTMPAVRTFALTGLAVILDF 784
Db 427 MSPROASESMRELFGANIATSLVMVFPVPAFFPG--TTGQIYKQFALTIAFSWVIST 484
Qy 785 LIQMSAFVALLSLDSKRGQASRLDVCCCVKQPELPPGQEGELLGFFQKAYAPP--LLH 842
Db 485 FLALTLPSELALLRRGQRP-----GWLKGIHQVNRFLD 521
Qy 843 WITRG-----VLLLELALF-----G 858
Db 522 WMRRGYERSLNFKVRAIVLLEFLASIALTGWVLSVPTAFIPEDQGYFITIQGPEG 581
Qy 859 VLSYSMCHISVGLDOB-LALPK--DSYLDYFLNRYFEVGAPYVFTTILGYNFSSRAG 915
Db 582 VSLNYSNVMAQVEKILKLPVETGTFAGIGFSFGSNSANG--VIFTLPQWDERQPG 639
Qy 916 MNA---ICSSAGC-----NNFSF----- 930
Db 640 QSAPEIIGLAGVFSNITEARIFVSPPPIDGLSGFGQFPELODRAGNTGLNLLQVMG 699
Qy 931 -----TQKIQ-----YATEPPE-----QSYLAIPASSW 953
Db 700 EIMRGNQTPGLOGVFSTFSANTPQMLIDIRNKAKALQVDVDEVENTLQSVL-----GSRV 756
Qy 954 VDDFDWLTSPSSCCRLYISGPNKDKFCPSTVNSLN-----CLKNCMSITMGSVRPS 1004
Db 757 VDNF---NQOQTYRYVYQADAQFRSNPEDIGKLYVRSANDQMISLSNLVTITPATGAQT 813
Qy 1005 VEQFHKYLFWLNDNRNICKPKGGLAA--YSTSVNLTSQGVLASRPMAYHKPLKNSQDY 1062
Db 814 INHYNLFRSITIN-----GAAAPGYSS-----GOATTAMEKLAKEVLPAIMGY 856
Qy 1063 TEALRAARELANITADLRKVPCTDPAF---EVFPYITITNVFYEQYLTILPEGLFMLS 1118
Db 857 E-----WSGITAEESGGQAPLIIFGLGLVFLVLAQYENYVDPL---IIMLSV 904
Qy 1119 CLVPTFAVSC-----LLGL-----DLRGLNLISVM- 1147
Db 905 FLAILGALSQAQLRGLSNDVFCQVGLVMLIGLAKNAILIYEFANQLREGGLPIIKATVQ 964
Qy 1148 -----ILVDTVGP-MALWDISYNVSLINLVSAVGSVEPVSHITRSPALSTKPTWL 1198
Db 965 ASQERLRPIMLTSLSLTGLVM-----PLINPIGA-----GAASRK----- 999
Qy 1199 ERAKEATISGSAVFAVAGVAMTNLPGLVL 1227
Db 1000 -----SLGTAVVGMIVSTVLSLVV 1020
```

RESULT 52
T20342

hypothetical protein D2013.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20342

R:Mortimore, B.
submitted to the EMBL Data Library, January 1995

A:Reference number: Z19259

A:Accession: T20342

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1032 <MIL>

A:Cross-references: UNIPROT:Q18968; UNIPARC:UPI00001642EB; EMBL:Z47808; PIDN:CAA87777.1;

A:Experimental source: clone D2013

C:Genetics:

A:Gene: CESP:D2013.8

A:Map position: 2

A:Introns: 25/2; 73/1; 217/1; 248/3; 570/2; 695/3; 775/1; 807/3; 953/3; 995/3

Query Match 2.0%; Score 136.5; DB 2; Length 1032;

Best Local Similarity 19.9%; Pred. No. 0.31;

Matches 100; Conservative 81; Mismatches 199; Indels 123; Gaps 20;

Qy 337 LGQFFQGWGTWASWPLTILVSVIPVVALAAGLVFTETDTPVELMSAPNSQARSKAF 396

Db 16 IGNAVHDYGRLCRAHPKACLSMLLTWILVSYPTI-TRLRL-PV---STPIDVFWSEHLH 70

Qy 397 HDQHFGFPFERTNQVILTAPNRSSVRYD---SLLLGPKNFSGIILDLLELLELLEQLERLH 453

Db 71 VNDKIAPFW-INE-----NPASYIQOFIVSTTISPNNATEMGEHAVRAAIATAFRIRQ 123

Qy 454 LQWSPAEQARNISLQICYAPLNPDTNSLYDCCINSL-----LQYFQNNRTL 500

Db 124 ILLAEP-----AVEELCLRLANQRQDSSWPFRRSKLCVLSPASIWYNNLQKFEDEDDT 177

Qy 501 LLLTANQTLMGQTSQVDWKHFLYCANAPLTFKDGDTALALSCMD--YGAPVFPFLAIGG 558

Db 178 ITTVFNE-----HCKS---TF-----CMRDLILGAPI-----AATG 205

Qy 559 YKGKDYSAEALII---MTFSLNNYPAGDRLAQAKLWEEAFLEEMRAFQRMAGMFQVTF 615

Db 206 IKQYQYTRKKKIEFAVTMFFARYSKKVIQGIKELQKEFLVDTPNDQRT--FVQYVF 263

Qy 616 TAERSLEDEINRTTAEDLPATSYIVIFLYISALGSYSWSRWVDSKATLGLGGVAV 675

Db 264 HPLKTFSDYI-----FLIST-YFVCMIVV-----YYSRKKIQMWASRWGLAFASSFT 309

Qy 676 VLGVAMAMGFFSYLGTRSSVLIVQVPPFLVLSVGADNIFIFVLEYQRLPRRPGPREVH 735

Db 310 VASTLLMTGTICAHLDLSLTTTGWSEVVPYIALIMGLNTLC----- 350

Qy 736 IGRALGRVAPSMLLCS-----LSEALCFELGALTMPAVRTFALTSGL 778

Db 351 ITRSVVYTSPLDVSRRHAHGLSQEGYKLTQYVILELLALLIGLFTTRISDIQEQCFQSVI 410

Qy 779 AVILDFLLQMSAFVALLSLSKR 801

Db 411 CVTVDFYMQLPFYAPCLTFDLQR 433

RESULT 53

G81825

probable drug efflux protein NMA1969 [imported] - *Neisseria meningitidis* (strain Z2491)

C:Species: *Neisseria meningitidis*

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C:Accession: G81825

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,

; Holroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: G81825

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1067 <PAR>
A:CROSS-references: UNIPROT:Q9J7S1; UNIPARC:UPI00000C4CBF; GB:AL162757; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: mtrD; NMB1969
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 135; DB 2; Length 1067;
Best Local Similarity 20.9%; Pred. No. 0.41;
Matches 109; Conservative 83; Mismatches 199; Indels 130; Gaps 25;

QY 452 RHLQVW-SPEAQR--NISLDICYPALNPNDNTSLYDCCINSLLQYFQNNRTLLLTANOT 508
DB 180 RAMRIWVDPKKLQNTNLSFADVGSA--LSAQNVOISAGSIGSLPAV--RQGT---VTATVT 233
QY 509 LMGQTSQVMDKDFLYCANAPLTFKDGTTALALSCMADYCAPVFPFLAIGGYGKDYSEAE 568
DB 234 AQQLGTAEFEFGNVILRANT-----DGSNVYLKDVAVRGLGM-----EDYSST 277
QY 569 ALI-----MTPSLNNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMPQVTFTAERSLE 622
DB 278 RLNGVNTTGMVNLN--SGN-AMATAKAVKERMATLEKYFPQGMS--WKTPTDTSKFVE 332
QY 623 DSNRTTAEDELPIFATSYIVIFLYISLALSGSYSSSRVMDSKATILGLGGVAVLGA VMA 682
DB 333 ISIEKVIHTLEAMVLVFMVYLFQNI--RYTLPTIVVPISL-----LGGFA----- 379
QY 683 AMGFPSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYORLPRRPGPREVHIGRALGR 742
DB 380 ---FISYMG--SINVLTWFMVVLVIGIVDDAIVVENVVERIMAGEGLPPKEATKKAMQ 435
QY 743 VAPSM-----LCSLSEALCPFLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSL 797
DB 436 ISGAVIGITAVLISVFPVPLAMFSGATNI--YKQFALTWASSI-----AFSAFLAL--- 484
QY 798 DSKRQBSRLDVCCCKVQDELPPQGGEL--LLGFFQKAYAPP-----LLHWIT 845
DB 485 ---TLTPALCATMLKTIKPHHEKKGFGFNKFNKNSWTHGYEGRVAKVLRKTF 536
QY 846 RGVLLFLALFGVSLYSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYFVT 905
DB 537 RMVVYIGLAVGV-----FLFMR-----LPTSFLT 563
QY 906 LGYNFSSEAGMNAICSSAGC-----NNFSFTQKIQVATEPPE 942
DB 564 EDQGFV---MVSVOLPAGATQERTNATLAQVTLAKSIPE 600

RESULT 54
E81051
multiple transferable resistance system protein MtrD NMB1715 [imported] - Neisseria men
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: E81051
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: E81051
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1067 <TET>
A:CROSS-references: UNIPROT:Q9JY67; UNIPARC:UPI00000C4798; GB:AE002521; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1715
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 135; DB 2; Length 1067;
Best Local Similarity 20.9%; Pred. No. 0.41;

Matches 109; Conservative 83; Mismatches 199; Indels 130; Gaps 25;

QY 452 RHLQVW-SPEAQR--NISLDICYPALNPNDNTSLYDCCINSLLQYFQNNRTLLLTANOT 508
DB 180 RAMRIWVDPKKLQNTNLSFADVGSA--LSAQNVOISAGSIGSLPAV--RQGT---VTATVT 233
QY 509 LMGQTSQVMDKDFLYCANAPLTFKDGTTALALSCMADYCAPVFPFLAIGGYGKDYSEAE 568
DB 234 AQQLGTAEFEFGNVILRANT-----DGSNVYLKDVAVRGLGM-----EDYSST 277
QY 569 ALI-----MTPSLNNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMPQVTFTAERSLE 622
DB 278 RLNGVNTTGMVNLN--SGN-AMATAKAVKERMATLEKYFPQGMS--WKTPTDTSKFVE 332
QY 623 DSNRTTAEDELPIFATSYIVIFLYISLALSGSYSSSRVMDSKATILGLGGVAVLGA VMA 682
DB 333 ISIEKVIHTLEAMVLVFMVYLFQNI--RYTLPTIVVPISL-----LGGFA----- 379
QY 683 AMGFPSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYORLPRRPGPREVHIGRALGR 742
DB 380 ---FISYMG--SINVLTWFMVVLVIGIVDDAIVVENVVERIMAGEGLPPKEATKKAMQ 435
QY 743 VAPSM-----LCSLSEALCPFLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSL 797
DB 436 ISGAVIGITAVLISVFPVPLAMFSGATNI--YKQFALTWASSI-----AFSAFLAL--- 484
QY 798 DSKRQBSRLDVCCCKVQDELPPQGGEL--LLGFFQKAYAPP-----LLHWIT 845
DB 485 ---TLTPALCATMLKTIKPHHEKKGFGFNKFNKNSWTHGYEGRVAKVLRKTF 536
QY 846 RGVLLFLALFGVSLYSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYFVT 905
DB 537 RMVVYIGLAVGV-----FLFMR-----LPTSFLT 563
QY 906 LGYNFSSEAGMNAICSSAGC-----NNFSFTQKIQVATEPPE 942
DB 564 EDQGFV---MVSVOLPAGATQERTNATLAQVTLAKSIPE 600

RESULT 55
H91170
hypothetical membrane protein [imported] - Escherichia coli (strain O157:H7, substrain :
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H91170
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H91170
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-749 <HAY>
A:CROSS-references: UNIPROT:Q8X6L3; UNIPARC:UPI0000165513; GB:BA000007; PIDN:BA037759.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC84336

Query Match 1.9%; Score 134; DB 2; Length 749;
Best Local Similarity 19.0%; Pred. No. 0.3;
Matches 162; Conservative 129; Mismatches 313; Indels 248; Gaps 40;

QY 455 QWSPEAQRNISLDICYPALNPNDNTSLYDCCINSLLQYFQNNRTLLLTANOTLMQ 512
DB 62 QEWLTLLQKSAALGDV--KGPMDASQQAQWGA-----FFWQHRNGLIDPNTRARLONG 113
QY 513 TSQVDWKDHFYLYCA-----NAPLTKDGTALALSCMADYCAPV----- 550
DB 114 EAQAWILSOLYSAPSGVSGKELQNDPLMLRGSQA---MAKQQRLLMDGWLVTQDP 170
QY 551 ---FPFLAIGGYGK--DYSEAEALIMTFSL-----NNYPAGDPRLAQAKLWBEAFLE 599

Db	171	QGNWYLLHGLAGSDFMQOOTHOLIITLNTLEKDLKTRYP-----QAQLLSR-----	218
Qy	600	MRAFORRMAGMFOVTTFAERSLEDEINRTTAEDLPFIATSYIVIFLYISIALGSSWSR	659
Db	219	-----GTVFYSDYASQAQKD-----ISTLGVATLLGVILLIIVAVFRSLR	258
Qy	660	VMDVKATLGGVAVVLGAVMAAMGPFYSIGIRSSLVILQVFLVLSVGADNIPFVL	719
Db	259	PLLLCVISIGICALA---GTVATLLIFGEHLMTLVMSMSVI-----GISAD-----YTL	305
Qy	720	EYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGGALTPMPAVRTFALTSGLA	779
Db	306	YI--LTERWVGNDSVPQWSLAKVNALLLALLTTVAAYLLIMLAPFGIR-----	354
Qy	780	VILDFLLQMSAFVAL--LSLDSKROBPASRLDYC-----CCVKPQBLPPPGQEGILLGFFQ	833
Db	355	-----QMAIFAAVGLS-----ASCLTVLFWHPMLCRGLPVRVPFAMA-----	391
Qy	834	KAYAPFLHWHITR-----GVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDY	886
Db	392	-----LMLRWLAAMRRNKKLSLGLPVALALFSLAGMSMLRVDDDISQLALPQ-----H	440
Qy	887	FLFLNRYE--VGAPV---YFVTTILGYNFSSBAGNNAICSSAGCNPFSTQKIOVA--TEF	940
Db	441	ILAQEKAITALTGQSVQDKWFV---YGDSPOQTLRRL-----EKYTSLEVAKEG	489
Qy	941	PRQSYLAIPASSWDDFDIMLTPSSCCRLYISGPNKDKPCPST--VNSLNCILKNCMSITMG	999
Db	490	LISNRYTTPNLNLAQEEDLQL-----LKTAAPTVTKALQNAGLTAVNPDLNAMPVND	543
Qy	1000	S--VRPSVEQFKHYLPWFL--NDRPNIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKP	1055
Db	544	EWLASPASGEWR--LLMLTLENAGESGVLPVVEG-----VKSALM-----	581
Qy	1056	LKNSQDYTEALRAARELANIATADLRKVPGT---DPAEVFPYITINVFYEQYLILPEG	1112
Db	582	-----QEIATYPCGTAWVDRKSTFDELFALRYVLT-----G	614
Qy	1113	LFMLSLCLVPTFAVSCLLGLDLRGLANLLSIVMLVDTVGFMAWDISYNAVSLINLV	1172
Db	615	LLVALAVIACGAVA---RLGWKGLISLPSVLSLGCGLAVLAMSQAVNLFSLLALV	670
Qy	1173	SAVMSVVEFY-----SHITRSPAISTKPTWLERAKEATISMGSAVFAGV--AMTNLPG	1223
Db	671	LVLVIGIGINTLPFSNPRGTPFLTSLALAL-----AMLTLLTLGLMLVFSATQALSSRGI	724
Qy	1224	ILVLGLAKAQLI	1235
Db	725	VLVSGIFTAFELL	736
RESULT 56			
D69403			
conserved hypothetical protein AF1229 - Archaeoglobus fulgidus			
C;Species: Archaeoglobus fulgidus			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004			
C;Accession: D69403			
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson			
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.			
.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.			
Nature 390, 364-370, 1997			
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.			
Smith, H.O.; Woese, C.R.; Venter, J.C.			
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae			
A;Reference number: A69250; MUID:98049343; PMID:9389475			
A;Accession: D69403			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-750 <LSE>			
A;Cross-references: UNIPROT:O29039; UNIPARC:UPI0000056DE1; GB:AE001019; GB:AE000782; NIT			
C;Superfamily: Archaeoglobus conserved hypothetical protein AF0459			
Query Match 1.9%; Score 133; DB 1; Length 750;			

Best Local Similarity 19.5%; Pred.No.0.36;			
Matches 137; Conservative 117; Mismatches 273; Indels 176; Gaps 33;			
Qy	597	LEMRAPORMAGMFOVTTFAERSLEDEINRTTAEDLPFIATSYIVIFLYISIALGSSYS	656
Db	166	VEKVISITNPPHGLV--LQLTGGPALSYQIERETKSGFI--TMMVSILMLVLLFLTFSG	222
Qy	657	WSRVVDSKATLGLGGVAVVLGAVMAAMGPFYSIGIRSSLVILQVVPFLV---LSVGADN	713
Db	223	-----AVRKKITAFPLVISVM--SVTVVGLMPILGIPLSEHTNGALPMLIGLAIEYGAQ-	276
Qy	714	IFIEVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGGALTP--MPAVRTF	772
Db	277	---LQNRYE--ERREGRDVDAVVISITRTGLAIVM--ALITTVIGFMSMLAPGMPAMAQF	331
Qy	773	ALTSGLAVIDFLQMSAFVALLSL-----DSKROEASRLDVCCVKPQELPPPGQEG	827
Db	332	GIISLGLIVAYLLTTLTFLPAVLKLDHWNDAKEQK-----REBEK	372
Qy	828	LLGFFQKAYAPFLHWHITRQVLLL---FLALFGVSLYSMCHISVGLDQELALPKDSYLL	884
Db	373	SVGTLNLSLSAISTLTATRPVGLVVAISIIVLFG--LYAAPQIGLETNKNYKVPQDLTAM	430
Qy	885	DYFLFLNRYEFGVAPYVFTTLGYNFSSBAGNNAICSSAGCNPFSTQKIQATEFFPQS	944
Db	431	QRFKEIERL--VGGQATYTLVLEVD-----ELNA-----ETLKEIDELS	467
Qy	945	YLAIPASSWDDFDIMLTPSSCCEL-----YISGPNKDKPCPSTVNSLNCILKNCMS	995
Db	468	KYIIVEKEELIYD-----SSITKLISEVRKAVGFEGLPESDAELMQILSAL-----	513
Qy	996	ITWGSVRPSVEQFKHYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKP	1055
Db	514	-----PQEVSRVI-----SGGQIAVHFPSSNADSQDE-----YISTHK-	546
Qy	1056	LKNSQDYTEALRAARELANIATADLRKVPGTDPAPFVP-----PYTIN--VPYEQYLIT	1109
Db	547	-----SIVRDV-----EFFGMSGYVTVGPPVIYGMGRIM	577
Qy	1110	PEGLPMLSLCLVPTFAVSCLLGL--DLRGLNLLSTVMIL--VDTVGFV-----ALWD	1160
Db	578	TSG--QTMTVVAVALIIMLLAVRSIRKAVFLLIATSIVIGAMNIMFLTGTQWTIS	635
Qy	1161	ISYNAVSLINLVSAVMSVFEVSHI--TRSPAISTKPTWLERAKEATISMGSAVFAGVAMT	1219
Db	636	IAMSITL-----GLGIDFSIHVLERYFEERAAPSPVEAVRRTIERTKAITTS--ALT	687
Qy	1220	NLPGILVLGLAKAQLIQIFPFRNLNLLITLLGLLHGLVFLPVIL	1262
Db	688	MAGGPGSLMFSTFPIMQNGF--IALVAILFSLAALTVPVPAPL	729
RESULT 57			
C69307			
conserved hypothetical protein AF0459 - Archaeoglobus fulgidus			
C;Species: Archaeoglobus fulgidus			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004			
C;Accession: C69307			
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson			
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.			
.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.			
Nature 390, 364-370, 1997			
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.			
Smith, H.O.; Woese, C.R.; Venter, J.C.			
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae			
A;Reference number: A69250; MUID:98049343; PMID:9389475			
A;Accession: C69307			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-736 <LSE>			
A;Cross-references: UNIPROT:O29790; UNIPARC:UPI00000570DE; GB:AE001073; GB:AE000782; NIT			
C;Superfamily: Archaeoglobus conserved hypothetical protein AF0459			

A;Title: Structural organization of the conserved gene block of Herpesvirus saimiri codi
A;Reference number: A34126; MUID:90163221; PMID:2154888
A;Accession: D34126
A;Molecule type: DNA
A;Residues: 1-808 <ALB>
A;Cross-references: UNIPARC:UPI0000138691; GB:M31122; NID:g331052; PIDN:AAA46164.1; PID:
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-808/Product: glycoprotein B #status predicted <GLB>
F;675-692/Domain: transmembrane #status predicted <TM1>
F;702-722/Domain: transmembrane #status predicted <TM2>
F;730,158,239,251,285,331,344,355,361,471,532,569,587,598,727/Binding site: carbohydrate
Query Match 1.9%; Score 130.5; DB 1; Length 808;
Best Local Similarity 21.3%; Pred. No. 0.59;
Matches 92; Conservative 71; Mismatches 162; Indels 107; Gaps 18;
QY 333 TSLSDKLS--PSTHTLLG--QFFQGWGTWVASWPLTILVLSVIPVVALAAGLVFTLTD 378
DB 345 TTLEDKIARVNNHTPNGTAETAYOTEGGMILVWQ-----PLIAELESEAMLEATTS 395
QY 379 PVLEWSAPNSQARSKAFHQHFGFPFRTNQVILTAPNRSRYRDSLLGPKNFSGLIDL 438
DB 396 PVPF-SAPTSSRSKRAIRSIRDSAGSENNVFL---SQIQAYDKLRQSINN----- 444
QY 439 DLLLELLEL---OERLRHLQWSPQAQRN-ISLQDICYA-PLN-----PDNTSLVDCCINS 489
DB 445 --VLEELATWCREQVRQTMWYEIAKINPTSVMTAAYGKPSRKALGALVISVTEC---- 498
QY 490 LIQYQNNRTLLLTANOTLMQTSQVDWKDHFLYCANAPLTFKD----- 534
DB 499 -----INVDQSSVSIHKSLENTENDICYSRPPVTFKFNSSQLFKGQLGARN 545
QY 535 ----GTALALSQADYGAVPFFLAIGGYKGKDYSEAE-----ALIMTF-- 574
DB 546 EILLSESLENC--HQNAETFTTAKNETYHFKNYHVETLPVNNISTLDTFLALNLTIE 603
QY 575 -----SLNNYPAGDPPLAQAQKWEAEFLERAFORRMAGMPOVTTFAERSLEDEINKTT 629
DB 604 NIDFKAVELYSSGERKLANVDFETWTF-REYNYYAQSIGSLGRKOPDNSQRNNRRIIQDF 662
QY 630 AEDLPFATSYIVIFLYISLALGSYSSWSRVMDSKATLGLGGVAVLGAVMMAAGFFSY 689
DB 663 SEILADLGSIGKVI---VNVASGAFSLFGIVT-----GILNFKNPLGGMFTLL 710
QY 690 LGIRSSSLVILQV 701
DB 711 IGAVIILVILV 722
RESULT 60
A82357
probable multidrug resistance protein VC0164 [imported] - Vibrio cholerae (strain N16961
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82357
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Frazer, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82357
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1036 <HEI>
A;Cross-references: UNIPROT:Q9KVI2; UNIPARC:UPI00000C2BFE; GB:AE004106; GB:AE003852; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0164
A;Map position: 1
C;Superfamily: acriflavin resistance protein

Query Match 1.9%; Score 130.5; DB 2; Length 1036;
Best Local Similarity 18.0%; Pred. No. 0.82;
Matches 219; Conservative 180; Mismatches 399; Indels 419; Gaps 58;
QY 285 LVLIILICSVF-----AVVTI-----LLVGRVAPADKSKMVDPKKG 322
DB 20 LIVLLGLQAFKMQVREYPMNTNVTITVSTGYGASADLIQTQPL-----EQA 70
QY 323 TSLSDKLSST-HTLLQPFQFGHGTWVASWPLTILVLSVIPVVALAAGLVFTLTD 381
DB 71 VAQADNIDFMTSQSLGR-----STITVT-----MKLNTDPNA 103
QY 382 LWS---APNSQARSE--KAFHQ-----HFGPFPRTNQVILTAPNRSRYR 423
DB 104 ALSDVLAKTNSVRSQLPKEAEDFTVMTSGTSTTAVLYIG--FTSNEL-----ASSQITD 155
QY 424 SL--LLGPKNF--SGILDLDLLELELQERLRH-LQVMSPEAQR---NISLQDICYAPL 475
DB 156 YLERVVPOLYAVNGVSSID-----LYGMYALRVWLDPAKMAALNSAADV-MSIL 207
QY 476 NPNTS-----LYDCCINSLLQYFQNNRTLLLTANOTL--MGQTSQVDM-KDH 521
DB 208 NANNYQSATQATGEFVLYNGSADTQVSTVEDLESVVKAEGTIVTFLGDIKVLAKSH 267
QY 522 FLYCANAPLTFKDGITALALSQADYGAVPFFLAIGGYKGKDYSEAEALIMTFSLNYP 581
DB 268 DTYRASA-----KDVLEMLPELQKNMPSNIEMNVLYDSTV--AINESIHEVIKTI 289
QY 582 GDP-RLAQAQKWEAEFLERAFORRMAGMPOVTTFAERSLEDEINNTTAE 640
DB 290 ANPINIA-----KDVLEMLPELQKNMPSNIEMNVLYDSTV--AINESIHEVIKTI 341
QY 641 IVIFLYISLALGSYSSWSRVMDSKATLGLGGVAVLGAVMMAAGFFSYLGIRSS 700
DB 342 LIVLVVITLFGSL---RAVLPIVITPLISLIGVAM--VMQMGFSWNL-----MTLLA 390
QY 701 VVPLFVLSVQADNIFIFVLEYQRLPRPGEP-REVHIGR---ALGRVAPSMLLCSLEAI 756
DB 391 MV--LAIGLVDDAI VLENVDRIHKEGESPFRAAIIGTREIAIPVIAMTLTLGAVY 448
QY 757 CFFIGALTPMPAVTTPALTSLGAVILDFLLQMSAFVALLSLDSKQASR-----L 807
DB 449 A-LMGGITG-SLFKEFALTLAGSVFVSGIVALTLSPMWCSQMLKANEPKFKVHLL 506
QY 808 DVCCVCKPQELPPPGQGEGLLGFQKAYA--PFLHMI-----TRGVVLLFLALF 857
DB 507 DRMTARYERMLTAVMAHRPVVIAFAIVFASLPMLFKFIPSELAPSEDKGVIMLMTG 566
QY 858 GVSLSYMSCHISVGLDQELALPKDSVLLDYFLFL---NRYPEV----- 896
DB 567 NANLDYLANTMDDVKNILSDQPEVQFAQVFTGVFNSNOAFGIASMPWPSOREASQAT 626
QY 897 -----GAPVVFVTTILGYNFSS-----EAGMN 917
DB 627 RVGTLVQDIPGMVAVTAFOMPELPGAGSGLPVQFVITTFPSNFESLFTIATDVL 686
QY 918 AICSSAGCNFNSFTQIKQ-----YATEFPE-----QSYLAI 948
DB 687 FVYSDDLNFDSATWTKINIDKAGAGVVTMQDIGITLSTWMDGYVNRIDLNGRSYEVI 746
QY 949 P--ASSWDDDFIDWLTTPSSCCRLYISGNKDKFCP-----STVNSLCLNCKMSI 996
DB 747 PQVERKW-----RLNPESMNSYYRAAD-GKVIPLGSLVTIDVVAEPRLPHFQ 799
QY 997 TMGSYRPSVEQPHKYLPMFLDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKL 1056
DB 800 TVGAVPAGCTAMGDNDANNFEN-LASKLPGK----- 829
QY 1057 KNSQDYTEALRAARELAANITADLRKVPGTDPAFVFPYITNVFYEQVLTLPGLF 1116
DB 830 --YSHDYMGEAR-----QYVT---EG--- 845

```
QY 1117 SLCLVTPFAVSCLLGL-----DLRSGLNLLSIWMLIVDTVGPMALWDISYNAVSLI 1169
Db 846 -SALYATFGLALAIIFLVLAIQFESLRDPLVIMSVPLAICGALIALAWGTATMNIYSQV 904
QY 1170 NLVSAGVMSVE---FVSHITRSPAISTKPTWLERAKEAT-----ISMGSAVFAGVAMT 1219
Db 905 GLITLVGLTKHGILICEVAKBQLNKLSRIEAVNHAAKVRRLRPIMTMTAMIAIGL--- 961
QY 1220 NLPGLILVLGLAKAQLIQIPFRLNLIT---LGLLHGLVFLPVLISYVGPVNP-ALAL 1275
Db 962 -IPLMATGAGAQR---FSIGIVVSGLAIGTLFTLPLVPLVIYSYLAERKHKLPLPVFV 1015
QY 1276 EQK-----RAEAVAA 1286
Db 1016 EDKOLEKLARIDEAKA 1032

RESULT 61
H86016
hypothetical protein Z4861 [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H86016
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H86016
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-772 <STO>
A:Cross-references: UNIPROT:Q8X6L3; UNIPARC:UPI0000165968; GB:AE005174; NID:g12518149; E
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z4861

Query Match 1.9%; Score 130; DB 2; Length 772;
Best Local Similarity 18.8%; Pred. No. 0.6;
Matches 160; Conservative 128; Mismatches 317; Indels 246; Gaps 39;

QY 455 QWSPBAQRNISLQDICYAPLAPNDNTSLYDCCINSLLQYFQNNRTLL--LTANQTLMGQ 512
Db 85 QEWLTLLQSAALGDV-KGPMDAASQAQWGA-----FFWQHRNGLIDPNTARLQNGG 136
QY 513 TSQVDWKDHFYCA-----NAPLTFKDTALALSCMADYGAPV----- 550
Db 137 EAQAQWILSQLYSAFSGVSGKELQNDPLMLMRGSQA---MAKNGQRLMDGWLVTQDP 193
QY 551 ---PFPFLAIGGYKGK--DYSEAEALIMTFSL-----NNYPAGDPRLAQAKLWEAEFLRE 599
Db 194 QGNYWYLLHGLAGSSFDMMQOHTLITLNTLEKOLKTRYP-----QAQLLSR----- 241
QY 600 MRAFORMAGMQVTTAARSLEDEINRTTABDLPIFATSYIVIFVYISLALGSYSWSR 659
Db 242 -----GTVFYSDYASQAQKD-----ISTLGVATLIGVILLIVAVPRSLR 281
QY 660 VMVDSKATGLGCVAVVLGAVNAAMGFFSYGLTRSSVLQVLPVLVSVGADNIFIFVL 719
Db 282 PLLLCVISIGIGALA---GTVATLLIFGELHMLTLMVMSMVI-----GISAD-----YTL 328
QY 720 EYQRLPRRGPGEVHIGALGRVAPSMLLCSLSEACFFLGLATPMPAVRTFALTSLGLA 779
Db 329 YY--LTERVMHGNDVSPWQSLAKVRNALLLATTTVAAYLIMMLAPFGIR----- 377
QY 780 VILDFLLQMSAFVALLSLDSKQKQEARLDVC-----CCVKPQELPPPGGEGLLIGFFOK 834
Db 378 -----QMAIPAAV-----GLXASCLTVFWHPWLCRGLPVRPVPAMA----- 414
QY 835 AYAPFLHWTIR-----GVVLLLEFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYF 887
Db 415 ----LMLRWLAWRNKKLSLGLPALALFSLAGMSMLRVDLDDISQLALPQ-----HI 464
```

```
QY 888 LFLNRYFE--VGAPV---YFVTTLGVNFSSEAGMNAICSSAGCNNSFTQKIOYA-TEPP 941
Db 465 LAQEKAITALTQSDVKWFV---YGDSPQOTLRRL-----EKYTASLEYAKKEGL 513
QY 942 EQSYLAIPASSWDDPIDMLTPSSCCRLYISGPNKDKFCEPT-VNSLNCNKNCMSITMGS 1000
Db 514 ISNYRTIPUNSLAROEEDLQ-----LKTAAPTVTKALQNALGTAVNPDNLNAMPVNVDE 567
QY 1001 ---VRPSVEGFHKYLPWFL--NDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAVHKPL 1056
Db 568 WLASPASEGWR--LLWLTLENGESGVLPVEG-----VKSALM----- 604
QY 1057 KNSQDYTEALRAARELANITADLRKPGT---DPAFVFPYTTITNVFEYQYLTILPEGL 1113
Db 605 -----QEIATYPGCIAMVDRKSTFDELFALYRYVLT-----GL 638
QY 1114 FMLSCLVPTPAVSCLLGLDLRSGLNLLSIWMLIVDTVGPMALWDISYNAVSLINLVS 1173
Db 639 LLVALAXIACGAVA---RLGWRKGLISLPSVLSGCCGLAVLAMSQQAVNLFSLALVL 694
QY 1174 AVGMSVEFV-----SHITRSFAISTKPTWLERAKEATISMGSAVPAGV-AMTNLPGI 1224
Db 695 VLIGIGNYTLFSPNPRGTPLTSLALAL-----AMLTTLLTGLMLVFSATQAISSEFGIV 748
QY 1225 LVLGLAKAQLI 1235
Db 749 LVSGIFTAPLL 759

RESULT 62
A89862
Na+/H+ antiporter subunit [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: A89862
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89862
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-801 <KUR>
A:Cross-references: UNIPROT:P60674; UNIPROT:P60675; UNIPARC:UPI000000CAC59; GB:BA000018;
A:Experimental source: strain N315
C:Genetics:
A:Gene: mnhA
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

Query Match 1.9%; Score 129; DB 2; Length 801;
Best Local Similarity 18.7%; Pred. No. 0.74;
Matches 131; Conservative 99; Mismatches 202; Indels 268; Gaps 31;

QY 641 IVIFVYISLALGSYSWSRVM-----VDSKATLGLGCVAVVLGAVNAAMGFFSYGLI 692
Db 39 IVIFVYMLTKITKTSYSGNTVMKTLNWMHPGNNFPLYDLGLGLPSLLISGIG----- 91
QY 693 RSSLVILQVVPVLVSGADNIFIFVLEYQRLPRRGPGEVHIGALGRVAPSMMLCSL 752
Db 92 --SLVVLVYISGLYSKSEKGNFYCYLL----- 116
QY 753 SEACFFLGLATPMPAVRTFALTSLGLAVIL-DPFLQMSAFVALLSLDSKQKQEARLDVCC 811
Db 117 -----LPMGAM-----LGVLSDNIVILYLFWELTSPSS----- 145
QY 812 CVKPELPPPGGEGLLIGFFOKAVAPFLHWTITGWVLLLF--LALFGVLSYSMCHISV 869
Db 146 -----FLLISFWERQAS--IYGAQSLIITVFGGLSLIGIIL----- 182
QY 870 GLDQELALPKDSYLLDYFLFLNRYFEVGPVYFVTTLGVNFSSEAGMNAICSSAGCNNS 929
```

Db 183 -----LAIPQSFQIY--MIQHASEIQNSPFFIFAM-----ILIMIG-----A 219
Qy 930 FTQKIYATEPPEOSYLAIPASSWDDPIDWLTSSC-----CRLYISGNKKKFC 980
Db 220 FTKSAQF-----PFYIWLDPAMEAPTVPVSAYLHSATWVKAGLYLIARMTIFA 267
Qy 981 PS-----TVNSINCLKNCMSITMGSVRPSVQFHKYLPWFILNDRPNIKCPKGGGLAAVSTS 1035
Db 268 ASQGVWTVT-----LVGLITLFWASLNATKQODUK-----GILAFSTV 306
Qy 1036 VNLTSDGOVLASRFMAYHKPKNSQDYTEALRAARELAANI-----1076
Db 307 SOLGIMAMWLGISYHYQGDDSKIYAAFTAAIFHLINHATFKGALFMITGAVDHSSTG 366
Qy 1077 TADLRKVPGTDPAPFVFTYTNVFEQYVILPEGLFMLSCLVPPFANVSCILLGLDLR 1136
Db 367 TRDVKKLGG-----LITIMPISFT-----ALSMAGVPPF-----400
Qy 1137 SGLNLLSTV-----MLVDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHIT 1186
Db 401 NGFLSKESLETFTTASQANLFSVDTLGTL-----PFIIGVSVFTFYVSIKFIHIL- 453
Qy 1187 RSFAISTKTPLW-ERAKEATISM--GSAVFAG--VAMTNLPGILVLGLAKAQLIQI-----1237
Db 454 --FFQYKPEQLPKKAHEVSIILMLSPAILATLIVILGLFPGILTNSIIEPATSSINHTV 511
Qy 1238 -----PF-----FRNLTLITLGLLHGLVFLPVILSY 1264
Db 512 IDDVBFHMFHGLTPAFLSTLVIYILGIL-----LIVTFYSY 546

RESULT 63
AB0370
multidrug efflux protein [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB0370
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001, MUID:21470413, PMID:11586360
A:Accession: AB0370
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1043 <KUR>
A:Cross-references: UNIPROT:Q8ZCB8; UNIPARC:UPI00000DCAD6; GB:AL590842; PIDN:CAC92285.1;
C:Genetics:
A:Gene: YPO3043
C:Superfamily: acriflavin resistance protein

Query Match 1.9%; Score 129; DB 2; Length 1043;
Best Local Similarity 18.8%; Pred. NO. 1.1;
Matches 185; Conservative 146; Mismatches 287; Indels 366; Gaps 54;

Qy 475 LNPNTSLVDCCNLSLQYFQNNRTLLLL-----TANQTLMGQTSQVDWKDHF 522
Db 188 LDPNKLNTYQLTTSIDVSAIQSQNTQVAVGQLGGTPAVDQNALNATINAQ-SOLOTPPEF 246
Qy 523 LYCANAPLTFK-----DGTALALSCMA--DYGAPVFPEL-----AIGGYKGDYSEAEALI 571
Db 247 -----REITLRVNQDGLSVLTGDAKIELGSEKYDYLSRFNGQAASGMGKILASGANEL- 300
Qy 572 MTFSLNYPAGDPRLAQAKLWEBAFLEEMRAF-----QRRMAGMFOVTTAERSLEDEIN 626
Db 301 -----QTDKRV-----KARLAELAPFPFHGLEAKIA--YETTPFVQASIKDVV- 341
Qy 627 RTTAEDLPFATYSYVIFELYISALGYSWSRWVMDSKATLGLGVA---VVLG--AVM 681
Db 342 KTLLE-----AILVFLVWYFLQNF-----RATL-IPTVAVPVVLLGTFVAVL 383
Qy 682 AAMGFFSYLGIRSLVILQVVPVFLVSVGADNIFIFVLEYQRLPRPG-EPREV---HIG 737

Db 384 SARGFSI-----NTLTWFAIV--LAIGLLVDDAIVVVENVERVMSEBGLDPREATRKSNG 436
Qy 738 RALGR-VAPSNMLCSISEALCPFLGALTMPAVRTFALTSLGLAVILDQLQMSAFVALLS 796
Db 437 QIQGALIGIALVSAVPIPMAPFGG--TTGAIYRQFSITIVSAMVL-----SVLVALIL 488
Qy 797 LDKRQESASRLDVCCCVKQBELPPPQGEGLLLGFFOKAY-----A 837
Db 489 TPA-----MCATLLKPIAFGHHAHAKRFGFWFNRMFDRNHRVYERVARVLHSLR 539
Qy 838 PFLHMTTRGVWLLFLAL-----FGVLSYMSCHISVGLDQELALPKDSYLLDYFL 888
Db 540 YMLLYLLLLGGLALLFLKLPSTPLPLEDRGVFM-AQVQLPVGSTQQOQLKVKVEKVENYFL 598
Qy 889 F--LNRVFEVGAIV-----YFVTILGY-----NFS-----911
Db 599 TEKKNVLSVFATVGSFGNGQNVARLFIRLADWDORTASTDSSPAIIERATKELSKIV 658
Qy 912 ----SEAGMNAICSSACNNFSF-----TQKIYATEFP-----941
Db 659 EAKVSVSPPAISGLGSSGDFMELQDHGGHDKLMVARNQLLQMASQEBPALTRVRHG 718
Qy 942 --EQSYLAIP-----ASSWDDPIDWLTFPSSCCRLYISGPN 975
Db 719 LDDSPQLQIDIDQKQAQALGVSLNDINSTLKTAMGSTYVNDVDV---RGRVKYVYVQSEA 775
Qy 976 KDKFCPTVNSLCLKNCMSITMGSVRPSVQFHKYLPWFILNDRPNIKCPKGGGLAAVST- 1034
Db 776 TARMLPEDVVK-----WYVRNK-----NGMVVPSAP 802
Qy 1035 SVNLTSDGOVLASRFMAYHKPKNSQDYTEALRAARELAANITADLRKVPQTDPAFEVFP 1094
Db 803 STTRWEYSGPRLRYNGY-----SALEIVGEAASGV-----TGTAMDVM- 842
Qy 1095 YTIITNVFEQYVLTLPGLFM-----LSCLVPT-FAVSCILLGLDLRSLGLNLL 1143
Db 843 -----EKLVSQLPNGFGLGEMTMSYQERLSGQAPALYAISSLVVFCL-AALYESW 893
Qy 1144 SI--VMLV--DTVGFM-ALW-----DISYNAVSLINLVSAGVMS-----VEFVSH 1184
Db 894 SIPFSVMLVPLVGVIGAVATWKGLENDVYFQ-----VGLTIIGLSAKNAILIVEPANE 949
Qy 1185 ITRSFATSKPTWLERAK---EATISMGSAVFAGVAMTNLP---GILVL-----1227
Db 950 LN-----NRGKDLVEATLEASRQRLRPLMTSLAFIFGLVLPMAISQAGSGSQ 997
Qy 1228 -----GLAKAQLQIIFPRL 1242
Db 998 HAVGTGVGMGMSISATVLAIFVPL 1021

RESULT 64
DB3393
RND multidrug efflux transporter PA2018 [imported] - Pseudomonas aeruginosa (strain PA01
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: DB3393
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: DB3393
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1045 <STO>
A:Cross-references: UNIPROT:Q9RG59; UNIPARC:UPI00000D420B; GB:AE004628; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2018
C:Superfamily: acriflavin resistance protein


```
QY 677 LGAVMAAGPFYSLGIRSSVLQVVPFLVLSVGADNIFIFVLEVRQRLPRPGPGPREVHI 736
Db 227 LATPFVFAVFA-MCKSLNTVMTTITPLLI-IGS-----AYGLFYNALFRFSKRE--- 277
QY 737 GRALGRVAPSMLLCSLSAICFFFLGALTPMPAVRTFAL--TSGLAVIDLDFLLQMSAFVAL 794
Db 278 --AVKHFKPILFMSLTTAAGFMSVFIDIRAFRELGLVSSGLAVVVLVIF----- 327
QY 795 LSLDSKQEARLDVCCCKFQELPPPCQGGGLLGFQKAYAPFLHWTIRG--VVLL 851
Db 328 -----TSGVEIFRNYTPKRT-----RSFGMKYVGKIKALIVLV 361
QY 852 LFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLPLNRYFVEGAPVYFVTTILGNFS 911
Db 362 VFLVMAALSPFLKRVQVGSMDVSYFERDSERKAYDLIVKFKNTREPIYLVLEKNVFFV 421
QY 912 SEAGMNAICSSAGCNCNFSFTQKIQYATEFPBQSVLAIPASSWVDDFDIWLTPSCCRUYI 971
Db 422 G-----TDSKILKELI 432
QY 972 SGPNDKFCPSTVNSLNCCKMSITMGSVRPSVEQFKHYLPWFPLNDPRNPKCPKGLAA 1031
Db 433 EKIEKSEYVSVVFPVDIPVPI-M-YTLSRTNP-----FLKTFVGDNRIR----- 476
QY 1032 YSTSNLTSQGVLASRFMAYHKPLKNSQDVTEALRAARELAANITADL---RKVPGTDP 1088
Db 477 --LIVNLTPGY-----EHVKKVVDLINEVSVSTGSHYVAGS-- 512
QY 1089 APEVFPYITNVFYBQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLSIVMI 1148
Db 513 -----VLINWDINESIMRSQISQSVLASILI---FAMVFII---FRPLTTLTSLVMIPI 559
QY 1149 LVDVTG----FMALWDISYNNVSLINLVSAVGMSEFVSHITRSPASTKPTWLERAKEAT 1205
Db 560 AFTTIVFNLFMALFGISLDVSTSIISGLIMGLVIDYSIHASEERLRDPLVWK----- 614
QY 1206 ISMGSAVFAGVAMTNLPGFLVLGLAKAQILQIFFFFLNLLITLLGLLHGLVFL----PVI 1261
Db 615 -NVGFSV-----LTNALG-LISGPAVLVLSSELALFNISLMLLIGVGAFTLLIVQPMI 667
QY 1262 L 1262
Db 668 L 668
```

RESULT 70
G87398
AcrB/AcrD/AcrF family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87398
R.;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11255647
A:Accession: G87398
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1047 <STO>
A:Cross-references: UNIPROT:Q9A8Z1; UNIPARC:UPI00000C72B9; GB:AE005673; NID:g13422529; H
C:Genetics:
A:Gene: CC1204

Query Match 1.8%; Score 126; DB 2; Length 1047;
Best Local Similarity 19.7%; Pred. No. 1.7;
Matches 198; Conservative 138; Mismatches 343; Indels 326; Gaps 50;
QY 500 LLLLTANQILMGQTSQVDWKDHFLYCANAPLTFKDGTFALALSCMADYAGAPVPPF----- 553
Db 137 IIIILNTSTTLNPLQADYADRYLV---ERMSTVDGVAQANL-----FGAKIYAMRIWLDA 189

RESULT 71
A75560
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: A75560
R.;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

```
QY 554 --LAIGYKGYKQYSEAEALIMTFSLN-----NYPAGD-----PRLAQA----- 589
Db 190 NEMAARGVTVDVESA-----LNAQNLLEPAGSLESAAKDFTIRVARSYATPBEFAR 241
QY 590 -----KLMBEAPLEKRAFORM--AGHFQVTF-----TAER 619
Db 242 LPLRAANSQSVYLRGLDGVARIEEGPDRRLFRNGIDQVIGLITRQSQANDVAISKAVR 301
QY 620 SLDEINRT-----TAEDL-PIFATSYIVIFLYISIALGYSYSWSVMV 662
Db 302 AEVEALNQTLPGTKKVIADVNSVFTAEAIHEVITWGLAIGLVVMVNLPLGSRWSALI 361
QY 663 DSKATLGLGGVAVVLGA--VMAAMGFYSYLGIRSSVLQVVPFLVLSVG--ADNIFIFV 718
Db 362 PS-----IVAPICILSTFIILPLGF-----SLNLLTLLA-LVLAVGLVWDDAIVVV 407
QY 719 LEVQRLPR--PQEPREVHIGRALGR-----VAPSMLLCSLSEAI CFFL GALTMPAVRT 771
Db 408 ---ENIQRRYDDGEPPLVAAERGTQVFFAVAVATTVVLSVFPAPLMPLPGTIGRLFVELA 464
QY 772 FALTSGLAVIDLDFLLQMSAFVALLSLDSKQEARLDVCCCKVQPELPFPQGGSEGLL--- 828
Db 465 VAIAAV-----AFSALLALSLSPMWSK-----LLRPAHGGSLARR 502
QY 829 ----LGPFQKAYAPFLH----W-ITRGVVLFLALFGVLSYMSCHISVGLDQELALPK 879
Db 503 VNKTMGALSASRASLEQLILGGWSAAASGVLLVLLAGFAAVLF-----VSLPKELVPAE 556
QY 880 DSYLLDFL-----FLNRYFEVGAPVYFVTTI-----GYNFSEAG 915
Db 557 DRGRVDVSI SAAEGAGYDYTSAIALKTEKVDKREAGVSERTIMTIPRFGGNSFNANG 616
QY 916 MNA-----ICSSAGCNCNFSFTQ-KIQYATEFPPEQ-----SYLAI 948
Db 617 VVALKPGERDKTAEVAEALNKELSRFTSVRAVASVRGPFQRCGGGGGGGTNVDPFIV 676
QY 949 PASSWVDDFI---DWLTP-----SSCRLVISGPNKDKFCPSTVNS 986
Db 677 G-----NDYVQLANWLKPI LAEAQDNPLARPRLDYEPTAPRLVQ-IDLDKAAATLGV-S 729
QY 987 LNCCLKMSITMGSVR-----PSVEQFKHYLPWFNL-----DRNPKCPKGLAAYS 1033
Db 730 AQSVGRALETMFGSSRRATTYIKSGQYDVLQTNLDQRRSIEDLNRNVRNTGAI VPLS 789
QY 1034 TSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPEVF 1093
Db 790 TVVTELRG-----DTPDRPRVDRLSRVTLTAQLN--FG----- 821
QY 1094 PYTITN--VFYEQVLTILP-----EGLFMLSCLVPTFAVSCILLGLDL 1135
Db 822 -YTVEDAVAFPKDQAACHPAPGVSVKWGGQAKYLEGGSGGIAI----AFGLALLLVFLVL 876
QY 1136 RSGLLNLL--SIVMLIVDTVGFMAWDI-----SYNAVSLINLVSAVGMS-----VEF 1181
Db 877 AAQFESWIHPAVIMLTVPALAGGLGFLGMLTGSTINTYTSQILIGIATKNGILIVVEF 936
QY 1182 VSHITRSPAISTKPTWLER--KEATISMGSAVFAGVAMTNLPGIL--VLGLAKAQILOI 1237
Db 937 ANQL-RDEGLSVREAVIEAALRLRPIIMTS--VSAAMGALPLMLWAGAGAGSRQTICA 992
QY 1238 FFFRLNLLITLLGLHGLVFLPVILSVYVGPDPVPALALEQKRAE 1282
Db 993 VIFTGAIPATLLITFIVPFVYDLLARFTKSPETWARQIEAVEAQE 1037
```

[illegible]

Db 318 --ELRKSDQAAVR-----TGV--GEFIRVLAELVIVLL-----VSFFSLG--LR 358
Qy 1137 SGLNLISIVMLVDVTGFMALWDISYNAVSLINLVSAGMSVDFVSHITRSPASTKPT 1196
Db 359 TGLVVALSIPVLAMTFAAMHVFYGLHKSIGLALVLALGLLVDIAIIVEMAVKMEQG 418
Qy 1197 WLERAKEATISGMSAVGAVMTNLPGLVLGLAKAQ-----LIQIFPFLNLLIT 1247
Db 419 Y-DRLKAASFMTSTAFPLMTGLTITAGFLPIATAOSGTGEYTRSLFQVVITIAL----- 472
Qy 1248 LLGLLHGLVFLPVLISYVGPVNPALPA 1274
Db 473 VVSWFAAVFVP-----YLGAKLPLDLA 495

RESULT 74
CB3260
hypoetical protein PA3079 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: CB3260
R:Stover, C.K.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: CB3260
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-793 <CO>
A:CROSS-references: UNIPROT:Q9HZC9; UNIPARC:UPI00000C5906; GB:AE004732; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3079

Query Match 1.8%; Score 123.5; DB 2; Length 793;
Best Local Similarity 21.9%; Pred. No. 1.8; Mismatches 211; Gaps 41;
Matches 154; Conservative 98

Qy 635 IPATSVIVIFLISALGSYSWSRVMDSKATLGLGVAVVLGAVMAAGFFSYLGIRS 694
Db 240 VMVVGFFGICLITLVL---LYWFTKCI--RTIAV--LITTLVAVLQGLNLVGFGL 292
Qy 695 SLVILQVDFPLVSV-----GANIFIFVLEYORLPRR-----PGEPR 733
Db 293 DPYSM-LVPFLIFAIQISHGVQKINGIALQSSGADN---ALMAARLTFRQLFLFG--- 343
Qy 734 VHIGALGRVAPSMLLCSLSEALCFGLALTPMPAVRTFAL--TSLGAVIL-DFLQMSA 790
Db 344 -----MIALADAVGFTLLVDIGVIRELAGASIGVAVIVFTNLILPV 389
Qy 791 FVALLSLDSKROEASRLDVCCVKQELPPQGBGLLIGFPQKAYAPFLHWTIRGVVL 850
Db 390 ATSYGISKAVQSKDQAV-----REHP-----FWRLSNFSPKVAP-----VSI 431
Qy 851 LFLFALFGVLSYMSCHISVG-IDQ-----ELALPKDSYLDY-FLFLNRYPEVGPVFTT 905
Db 432 ATALLMGGLVYGRHLKIGLDLQGAPELR-PDSRYLNDNDFIIRN----- 476
Qy 906 LGVNSFSEAGMNAI-CSSAGCNFSPTQKIQATEPPEQSYLAIPASSW-----VDDF 957
Db 477 --YSTSSDLVVMVWTKSPGECST-----HQAAMDDELAKWLENTEGVQSA 520
Qy 958 IDWLTSSCCRLYISGNPKDFCPSTVN-SLNLCKNMSITMG-----SVRPSVEQPHK 1010
Db 521 ISWTVS---KQVIKWMNEGKWLTSRNQDVLNNSIARADGLYNTDCSLAP----- 570
Qy 1011 YLPWFILNDPNKCPKGGAAVSTSVNLISDGOVLASRFMAVHKPLKNSQDYTEALRAAR 1070
Db 571 -LLVFLNDHK-----AETLDRAVAVAQFAENDKDLDLQFLAAGNAGI 613

Qy 1071 ELAANITADLRKVGTPDPAFEVFFYTTINVFYQYLITLPEGLFMLSCLVPTF-----AV 1126
Db 614 EAATN-----EVIKQSELIILVLVIYICV-----AAMCMI-IFRSPAAT 650
Qy 1127 SCULLGLDRLASGLNLISIVMLVDVTGFMALWDISYNAVSLINLVSAGMSVDFVSHI- 1185
Db 651 LCIVLPLILTSVLGNAL-----NAALGIGVKVATLPVIALGVGIGVDYGIYI 698
Qy 1186 TR-SFAISTKPTWLERAKEATI-SMGSAV-FAGVAMTNLPGLVLGLAKAQLIQI-PFF 1240
Db 699 TRUESFLRMGLP--LQAYYETLRSTGKAVLFTGLC-----LAIGVATWIFSARKFOA 749
Qy 1241 RLMLLIT---LLGLLHGLVFLPVLISYVGPVNPALAEQKRA 1280
Db 750 DMGLMLTFFMLLWNMFGALWLLPALRFL---INPA-KLQARKA 788

RESULT 75

E70826
probable membrane protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70826
R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70826
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-964 <CO>
A:CROSS-references: UNIPROT:O53784; UNIPARC:UPI000012F272; GB:AL021943; GB:AL123456; NID
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: mmpL5

Query Match 1.8%; Score 123.5; DB 2; Length 964;
Best Local Similarity 18.6%; Pred. No. 2.3; Mismatches 357; Gaps 48;
Matches 200; Conservative 140

Qy 348 VASPLITLVLVS-IPVVVLAAGLVFTLTDPVELMSAPNSQARSEKAFHDQHGPPFR 406
Db 36 ILGMLVITAVLVTVPQLETVGQIQAVMSPD-----AAPSMISM-----KHIGKVE 83
Qy 407 TNQVILTAPNRSSRYDSLLGLPKNFSGILDLDLLELLELLELLELRLHLOVMSPEAQRNIS 466
Db 84 -----EGDSDSAAMIVL-----EGQR--- 99
Qy 467 LDICYAPLNPDNTSLYDCCINSILQYFQNNRTLLLTANOTLMQTSQVDWKDHFLYCA 526
Db 100 -----PLGDAHAFYDQMTGRL-----QADTTHVQSLODFWGD----- 132
Qy 527 NAPLTFKDGDTALALSCMADYGAPVFPFLAIGYIKGKDYSE-----AEALIMTFSLN 577
Db 133 --PLT-----ATCAQSSDCKAAVQVKLAGNQGESLANESVEAVKTIIVERLAPPGVK 183
Qy 578 NYPAGDPPLAQAKLWEEAFLEEMRAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPFA 637
Db 184 VVVTGSAALVA-----DQQQAG-----DRSLQ-----VIEA 209
Qy 638 TSYIVIFLYISIALGSYSWSRVMDSKATLGL-----GGVAVVLGAVMAAGFFSYLGIR 693
Db 210 VTFTTIVMLLLVRSIIT-SAIML-TMVVLGLLATRGVAF-----LGFHRIIGLS 259
Qy 694 SSLVILQVVPFLVLSVGADNIFIFVLEYORLPRRPEPREVHIGRALGRVAPSMLLCSLS 753
Db 260 TFAITLLVV--LAIAAATDYAIFLIGRYOE-ARGLGQDRSAYYTMFGTAHVILGSGIT 316
Qy 754 EAICFFLGALTMPMPAVRTFA--LTSGLAVILDFLLQMS-AFVALLSLDSKROEASRLDVC 810

Db 317 IAGATFCLSPTRLPYFQTLGVPLAIGWIVVAAALTLGPALIAVTSRFGKLEPKRM---- 373
QY 811 CCVKQQLPPQGGELGLGFPQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVG 870
Db 374 -----ARVGRKVGAA--IVRW--PGPILGAVALLVGLTLPGYRTN 414
QY 871 LQELALPKDSYLLDYFLFLNRYFVGAPVYFVTLGYNFSSEAGHNAICSSAGCNSFP 930
Db 415 YNDRNYPADLPANEGYAAERHF-----SOARN----- 444
QY 931 TOKIOVATEFPQSYLAIPASSWVDDFIDWLPSSCCRLYISGPNKDKPCPSTVNSLNC 990
Db 445 -----PE--VLWESDHDMENSADFLVINKIAK-----AIFAVEGI 478
QY 991 KNCMSITMGSVRPSVEQFKYLPWFLNDRPNIKCPKGLAAYSTSVNLTS--GOVLASRF 1049
Db 479 SRVQAITRDPGKP-IE--HTSIPFL-----ISMGTSQLTEKNQDLTARM 522
QY 1050 MAYHPLKNSQDYTEALRAELANIT-----AD----- 1079
Db 523 LEQVNDIQSNIDOMERHSLTQOMADVTHEMVIOGMTGMVVDVEELRNHIADDFRPIR 582
QY 1080 -----LRKVPGTDPAPRVFPYTLNV--FVEQVLTILPEGLWLSLCL 1120
Db 583 SYFYWEKHCYDIPVCWLSRSVFTLDGIDVMTEDINNLPLMQRLDTLMPQTAMPEMI 642
QY 1121 VPTFAVSCLLGL--DLRSGLNLLSIVMLVDTVGFMAWDISYNAVS----- 1167
Db 643 QTMKSMKAQMLSHSTQEGLODMAQO--EDSAAWGEAFDASRNDSDSYLPEVPDNP 700
QY 1168 -----LINLVSAGVMSVEP-VSH-----ITRSPASTKPTWLERAKEA---TISMGS 1210
Db 701 FORGLEQFLSPDGHAVRFITISHEGDPMSQAGIARIKNT-----AAKEAIKGTPLSGS 754
QY 1211 AVFAG-----VAMTNLPGTLVLGLAKAQLIQFP----- 1239
Db 755 AYLGGTAMFDFKLDGNDYDMIAGISALCLIFIMLITRSVAAAVIGTVVLSLGA 814
QY 1240 --PRMLLI--TLGLH-LHGLV--FLPVLISYVGVDPNPALEOKRABEAAV 1287
Db 815 SPGLSVLIHQHILGIELHVLVLAIVIALLAVGADYN--LLLVARLKEEIHAGI 866
RESULT 76
G69154
conserved hypothetical protein MTH420 - Methanobacterium thermoautotrophicum (strain Del
C)Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G69154
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69154
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-676 <MTH>
A:Cross-references: UNIPROT:Q26520; UNIPARC:UPI000006649C; GB:AB000826; GB:AB000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH420
Query Match 1.8%; Score 122.5; DB 2; Length 676;
Best Local Similarity 20.4%; Pred. No. 1.7;
Matches 104; Conservative 65; Mismatches 160; Indels 181; Gaps 24;
QY 472 YAPLNPNTSLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPIT 531
Db 27 YQPLMYDMDSYN-----LRLASNLNEGELSEGWDRIYSYPPGVPL- 69
QY 532 FPDGTALALSCWADYCAPVFPFLATGKYGKDYSEAEALIMTFSLNYPAGDPRLAQAKL 591

Db 70 -----DY-PPLLPHLVTLY-----LLFTWLL-----PGLTDTAP 98
QY 592 WEEAFLEEMRAPQRMAGMQVQTTAERSLE--DEINRTTAEDLPFATSY----- 640
Db 99 WLPALIAP-----LAGV--AVFAARSLDCDDVSATTAGLLATAAPFPFRTTPGFPD 149
QY 641 -----IVIFYISLALGYSYSSRSMVMSKATLGL-----GGVAVVLGAVMAAM 684
Db 150 TDMFNVLPLAIIALLHLSLRDGDW-----RMAVLSGALMGVFAAAWNGQLISYIMALS 205
QY 685 GPPSVLGTIRSSVILQVVPFLVLSVGAONIFIPVLEYQELPRRCPGEPREHVHIGRA----- 739
Db 206 ILHSLLYRRRBLI-----FLITS-----TLIVTMDPRAIPSPVGLMKIPVARADPPD 256
QY 740 -----LGRVAPSMLLCSLSEACIFLGAITPM--PAVTEALTSGLAVI 781
Db 257 PYANITELQRFQDDVLMALGGLLAGLAGRSLFRWRDSNVLPVILWILTGNASLL 316
QY 782 -----LDFLQMSAFVALLSLDSKQBSRLDVCCKVQSELPPGQGGELLGF 831
Db 317 WGIREFSELLTAPLLITSALFLADLTASR-----VPP-----CF 350
QY 832 FQKAPAPELLHWITRGVLLFLALFGVLSYMSCHISVG--LDQ-----ELALPKDSYLL 884
Db 351 RRK-----LH-VAVAIMVLPSLIISTGQYSALHPRVDDGLLDAADYIRAKTPPTDVTVI 403
QY 885 D-----YPLFLNRYFEVGAPVVFVTTLGY 908
Db 404 SNWVGHFFAFMAR-----RPVNFDRLAY 428
RESULT 77
Q6811
ribonucleoside-diphosphate reductase (EC 1.17.4.1) large chain - human herpesvirus 4 (ser
N)Alternate names: ribonucleotide reductase large chain
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C:Accession: A03753; A22907; S32996
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A>Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713; PMID:6092825
A:Accession: A03753
A:Molecule type: DNA
A:Residues: 1-826 <BAM>
A:Cross-references: UNIPROT:P03190; UNIPARC:UPI000000C0CA; EMBL:V01555; NID:959074; PID:
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.;
Nature 310, 207-211, 1984
A>Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Contents: annotation; protein coding region
R:Gibson, T.; Stockwell, P.; Ginsburg, M.; Barrell, B.
Nucleic Acids Res. 12, 5087-5093, 1984
A>Title: Homology between two EBV early genes and HSV ribonucleotide reductase and 38K
A:Reference number: A22907; MUID:84247360; PMID:6330697
A:Contents: annotation
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase large chain
C:Keywords: deoxyribonucleotide biosynthesis; early protein; oxidoreductase; redox-activ
F187-403,825/Disulfide bonds: redox-active #status predicted
F387,391/Active site: Asn, Glu #status predicted
F389/Active site: Cys (cysteine thiol radical intermediate) #status predicted
Query Match 1.8%; Score 122; DB 1; Length 826;
Best Local Similarity 20.4%; Pred. No. 2.4;
Matches 96; Conservative 46; Mismatches 150; Indels 178; Gaps 26;
QY 68 GDHLLILKIKCPRLVGTNPQACCSAKOLVSLASL-----SITKALLTRCP-----ACSD 118
Db 229 GKHSLLMRMI-NSHVEYHNYGC---KRPVSVAAYMEPHWQIFKFLTKLPENHERCPG 284
QY 119 NFVNLHCHTSCPNQSLFINVTRVAQLG-----AGQLPAVVAVYBAFYQHSFAEQSY 169

Db 285 IFTGLFV-----PELFFKFLFRDTPMSDWYLFDPKDAGDL-----ERLYGEBFEREY 331

Qy 170 -----DSCSRVRPAAATLAVGTCGYGSA-----LCNAQRWLNFGQDGTGNGLAPL 216

Db 332 RLVTAGKCGRVSTKSLMFSTVN--CAVYKAGSPFILLKRCNAHFWRDLQGEANMA----- 385

Qy 217 DITFHLBPQAVGGIOPLENGVARCNESQGDVATCSCDDCAACPAIARPOALDSTF 276

Db 386 -----ANLCAEVLQPSKSVATCN-----LANICLPRLVAPLAVRAQADTQ- 429

Qy 277 YLQG-----MPG-----SLVLI--IILCSVFAVVTILLVGRVAPARDKS 314

Db 430 --GBELLALPRLSVTLFEGAGVGDGFSLARLRDATQCATFVACSILOGSPTYDSRDM 487

Qy 315 KM-----VDPKGTSLSDKL-----SFSHTLLG----- 338

Db 488 SMGLGVOGLADVFDLQGWYTD-PSRSLNKEIFEHMYFTALCTSSLTGLHTRKIFPGFK 546

Qy 339 --QFQGGTGWASVPLTILVLSVIPVALAAGLVFTLTDPPVBLWSAPNSQARSEKAF 396

Db 547 QSKYAGGFHW-HDW-----AG---TDLST-PRINSRLSERIVRDGLP 585

Qy 397 HDQH-----FGPFR-----TNQVILTAPNRSSVRY 422

Db 586 NSQFTALMPTSGCAQVTCSDAFYFAYANASTKVTNKEALRPNSFRWH 635

RESULT 78

T52483

hypothetical protein [imported] - Thermus aquaticus

C:Species: Thermus aquaticus

C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004

C:Accession: T52483

R:Fee, J.A.; Chen, Y.; Barquera, B.; Teasos, P.; Gennis, R.

A:Reference number: 226089

A:Accession: T52483

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-385 <PEB>

A:Cross-references: UNIPROT:Q9XK7; UNIPARC:UPI00000BC751; EMBL:L09121; PIDN:AAD38864.1

A:Experimental source: strain H88

C:Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MTH76

Query Match 1.8%; Score 121.5; DB 2; Length 385;

Best Local Similarity 23.5%; Pred. No. 0.94;

Matches 77; Conservative 36; Mismatches 119; Indels 95; Gaps 13;

Qy 669 GLGGVAVVLGAVMAAMGFYSYLGISSVLVQVPPFLVLSVGADNIF-----IFV 718

Db 22 GPGTIGLVEGVAEATASLFKVVGGRSLDRIGRRRPFLLILGYGLPALFRPLLALAQSPHV 81

Qy 719 LEYQRLPR-----RPGEPREVIHGRALG-RVAPSMLLCSLSBAICFFLQAL 763

Db 82 LLYRFLDRTGKRLTAPRDALTAESAPKEA-LGRAYGLHRSLDLTGLATGLPPLAFL---L 137

Qy 764 TPMPAVRFTLSGLAVILDPELLQMSAFVALLSLDSKQESRLDVCVCVPPQELPPGQ 823

Db 138 LPLLGLFRGVFWLSALPALLAFLLLFAV-----RETFR-----PPSPLFP--- 176

Qy 824 GGLLLGFFQKAYAPFLHWHITRGVVLLFLALFGVSLVSMCHISVGLDQELALPKDSYL 883

Db 177 ---LALHLSPGYRRF-----LLVSLGLFALALSSNAFLRL-L-KELGLSQBEVA 221

Qy 884 LDYFLF-----LNRYFEVGAPVYFTTLGYNFSSEA-----GMN 917

Db 222 LAYTLVNLLYALLAYPLGLADRVGLGRMVATGFGLYALVILGFAPWARTAFWALGFPLY 281

Qy 918 AICSSAGCNFSFTQIKIYATFEPQS 944

Db 282 ALYSAA-----FEGANRAVLATLVPEEA 304

RESULT 79

H91123

probable oxidoreductase ECs3960 [imported] - Escherichia coli (strain O157:H7, substrain C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: H91123

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H91123

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-477 <HAY>

A:Cross-references: UNIPROT:Q8XAL3; UNIPROT:Q8FDE8; UNIPARC:UPI00000D06C0; GB:BA000007;

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: ECs3960

Query Match 1.8%; Score 121.5; DB 2; Length 477;

Best Local Similarity 19.4%; Pred. No. 1.3;

Matches 90; Conservative 73; Mismatches 155; Indels 145; Gaps 23;

Qy 890 LNRFFEVG---APVVFVTTLCY-----NFSSSEAGMNA-ICSSAG----- 924

Db 29 INNIIEGLASAPMFFLATIFYFIPFCLIIAEFVSLNKNSEAGVYAVYKSSILGGRWAFIT 88

Qy 925 -----CNFSFT---QKIQYAT-EPPEQSYLAIPASSWDD-----FIDWLTSSCCR 968

Db 89 AYTVMFVNLFPFTSLLPRVIAVASVAFGLGYEYIMTPVATTIISMVLFASFSTWSTNGARM 148

Qy 969 LYISGPNKDKCPSTVNSLCLNCKMSITM-----GSVRP-----SVQFHKYLPW-FLN 1017

Db 149 L---GP-----ITSVTSTMLLLTSLYLLAGTALVGQVQADPTIVDAMIPNFWNAFLG 200

Qy 1018 DRPNIKCPKGLAAYSTSVNLTSDGQ-----VLASRFMAYHKPLKNSQDYTEALRAAR 1070

Db 201 VTTWIFMAGAESVAVYVNDVKGSKSFVKVILLAGIFIG-----VLYSVS 247

Qy 1071 ELAANITADLRKPGTDPAPFVF-----PYTITNVF----- 1101

Db 248 SVLINVFSSKELKFTGGSQVQFHGMAAYFGLPEALMNRVGLVSFTAMFGSLLMWTATP 307

Qy 1102 YEQVLTILPEGLFMLSCLVETFAVSCLLGLDLRSGLNLLSIV-MILVDTVGFMAWD 1160

Db 308 VKIPFSEIPEGIFGKK-----TVELNENGVPARAAWIOQLVIVLPIMLIGSNTVOD 360

Qy 1161 ISYNAVSLINLVSAVGMSEVF-----VSHITRSPFAISTKPTWLERAKEATISM 1208

Db 361 LMN---TIINMTAAASMLPPLFIMLAYLNLRAKLDHLPDRFMGSRRTGI-----IVVSM 412

Qy 1209 GSAVPA-GVAMTNLPGLILVLGAKAQLIQIPFFRLNLLITLLG 1250

Db 413 LIAIFAVGVASTFP-----TGANILTIIFVNVGGIVIFLG 448

RESULT 80

GS9598

probable oxidoreductase ygjI [imported] - Escherichia coli (strain O157:H7, substrain E1 C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: GS9598

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A95480; MUID:21074935; PMID:11206551

A:Accession: GS9598

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-477 <STO>

A:Cross-references: UNIPROT:Q8XAL3; UNIPROT:Q8FDE8; UNIPARC:UPI00000D06C0; GB:AE005174;

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: y9j1

Query Match 1.8%; Score 121.5; DB 2; Length 477;
Best Local Similarity 19.4%; Pred. No. 1.3; Indels 145; Gaps 23;
Matches 90; Conservative 73; Mismatches 155; Indels 145; Gaps 23;

QY 890 LNRYPEVG---APVYFVTTLGY-----NFSSEAGMNA-ICSSAG----- 924

DB 29 INNTELGLASAPMPFLATIFYFIPPCLLIAEFVSLNKNSEAGVYVAVKSSLLGRWAFIT 88

QY 925 -----CNNEST-----OKIQAT-EFPEQSYLAIAPASSWDD-----FIDMLTPSSCCR 968

DB 89 AVTYMFVNLFFPTSLLRVIAVASYAFLGYEYIMTPVATTIISVMVLPFAFSTWSTNGAKM 148

QY 969 LYSISPNKDKPCFSTVNSLNLCKNCSITM-----GSVRP-----SVEQFHKYLPW-PLN 1017

DB 149 L-----GP-----ITSVTSTMLLLTSLYLLAGTALVGGVQPADPTVDMINFNFWAFIG 200

QY 1018 DRPNIKCPKGLAAYSTSVNLTSDQ-----VLASRFMAYHKPLKNSQDYTEALRAAR 1070

DB 201 VTTWFMAGGAESVAVYVNDVKGSKSPVKVILLAGIFIG-----VLYSVS 247

QY 1071 ELAANITADLRKVGTDPAFEVF-----PYTITNVF----- 1101

DB 248 SVLINVFSSKELKFTGGSVQVFHGMAYFGLPEALMNRVGLVSTFAMFGSLLMTATP 307

QY 1102 YEOYLITLPEGLFMLSCLVPTFAVSCLLGLDLSGLNLISIV-MILVDTVGFMAWD 1160

DB 308 VKIFFSEIPEGIFGKK-----TVELNENGVPARAAMTQFLIVLPLMTIIPMLGNTVQD 360

QY 1161 ISYNAVSLINLSAVGMSVEF-----VSHITRSPFAISTKPTWLERAKEATISM 1208

DB 361 LMN-----TIINMTAAASMLPPLFIMLAYLNRAKLOHLPRDFMGSRRTGI-----IVVSM 412

QY 1209 GSAVFA-GVAMTNLPGILVLGLAKAQLOIQIFFRNLNLTLLG 1250

DB 413 LIAIFAVGVASTFP-----TGANILTIIFVNGGIVIFLG 448

RESULT 81

AE3557

acriflavin resistance protein d [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: AE3557

R:DelVecchio, V.G.; Kapratral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goltzman, E.; Selkov, E.; Eizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AE3557

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1024 <KUR>

A:Cross-references: UNIPROT:Q8YCY25; UNIPARC:UPI000005846D; GB:AE008918; PIDN:AAL53624.1;

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI10382

A:Map position: II

Query Match 1.8%; Score 121.5; DB 2; Length 1024;
Best Local Similarity 28.1%; Pred. No. 3.5; Indels 43; Gaps 12;
Matches 65; Conservative 36; Mismatches 87; Indels 43; Gaps 12;

QY 1120 LVPTFAVSLIGLRLSGLNLLSIVMLVDTVGFMAWDISYNAVSLINLSAVGMSV 1179

DB 346 VVIVLAVSVSLG---LRAGFVVSLSIPLVLAITFLSMLSDISLQSVSLGALIINALGLV 403

QY 1180 EBFVSHITRSPFAISTKPTWLER-----KEATISMGSVAF---AGVAMTNLPGILVLGLAKA 1232

DB 404 D-----DAMIAVEMVMVARLEHGDHPINKAATVYVSHTAFPLTGTTLV- IAGFIPIGLNS 457

QY 1233 QLQIQIFF-----FRNLNLT-----ILGLLHGLVFLPVLTSYVGPVDPNPALEOKRAE- 1281

DB 458 QAGEYTTFTLVVIAVSLVSWVAVLPAFLLGVTFLP-----KKMKPH---EEKRSRP 507

QY 1282 -BAVAAMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGAIS-NFLPNNGR 1330

DB 508 PEAFSRVLLLSMRHKWTTIITIVLVFI-----SVFGMGFIBRQFPQSDR 553

RESULT 82

AE0816

probable efflux pump STY2719 [imported] - Salmonella enterica subsp. enterica serovar Ty

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AE0816

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerston, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

, S.; Moulé, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Party, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AE0816

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1037 <PAR>

A:Cross-references: UNIPARC:UPI0000059ACA; GB:AL513382; PIDN:CAD07711.1; PID:gi16503697;

C:Genetics:

A:Gene: STY2719

C:Superfamily: acriflavin resistance protein

Query Match 1.8%; Score 121.5; DB 2; Length 1037;
Best Local Similarity 18.9%; Pred. No. 3.5; Indels 333; Gaps 47;
Matches 171; Conservative 127; Mismatches 274; Indels 333; Gaps 47;

QY 533 KDGTAALALSCMA--DYGAPVPFPLAIGYKGYDSEAEALIMTFSLNNYPA----- 581

DB 255 QDGSVKLDGVATVELGAEKYDLS-----RFGNVPASGLGVKLAS 295

QY 582 GDPRLAQAQKMEAFLEEMRAF---QRRMAGMFQVTTAERSLEDEINRTTAEDLPIPA 637

DB 296 GANEMATAKLVLDRLNELAQVPHGLEVKIA--YETTSFVKASIIDVVKTL---LEATA 349

QY 638 TSYIVIFLYISALGSSYSSSRVWDSKATLGGVAVVLGVAWMAAGFPFVGLGRSSLV 697

DB 350 LVFLVMYLFQ-----NFRATL-----IPTIAVPVVLMTGTSVLYAFGYSI 390

QY 698 ILQVVPFLVLSVG--ADNIFIFVLEYQRLPRPG-EPREV---HIGRALGR-VAPSMILC 750

DB 391 NTLTFAMVLAIGLLVDDAIVVENVERIMSEGLTPREATRKSQIQGALVGIAMVLS 450

QY 751 SLSEACFPGLGALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKRQSEASRLDVC 810

DB 451 AVFVPMAPFGG--TTGAIYRQFSITIVSAWLSVLVAMILTPAL-----C 493

QY 811 CCVFPQELPPGQGE-----GLLIGF-----FOKAYAPFL---LHWITRGVLL- 851

DB 494 ATL-----LKPLHKGEGHQGGFFGWFNRFTNRNAREYKGVAKILHRSLSRWLIYVLLIG 549

QY 852 ----LFLAL-----FGVSLYSWCHISVGLDQELALPKDSYLLDYFL-----F 889

DB 550 GMVLFLLRLPTSLPQEDRGMTTSI-----QLPGSGTQQQLTKVVKVENYFTTHEKNNI 605

QY 890 LNRIFYEVGA-----PVYFV-----TTLGYNF-----S 911

DB 606 MSVFTVSGSGGNGQNVARMFVRLKQWDARDPTTSGSFAIIERATKAPNQIKEARVPAS 665

QY 912 SEAGNNAICSSAGCN-----NFSFTQ-----KI 934

DB 666 SPPIAISGLSSAGFMELQDHAGHDALMAARDQLIBLAGKNSLSLTVRHNGLDDSPQL 725


```
QY 588 OAKLWEAFLEBEMRAFORRMAGMFOVTFPTAERSLEDEINRTTAAEDLPFATSYIVIFLYI 647
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 676 PS-VNE-----SKVSFPDLSQOMLSLSVIOCLK----- 705
QY 648 SLALGSYSRGRVMVDSKATGLGSGVAVVLGAVMAAMGFFSYLGIRSLVILQVVPFLVL 707
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 706 ---AGKKQW-----RTASLTNACAGL-----LAGLKVLPILVY 736
QY 708 -----SVGADNFIIFVLEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLEAI 756
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 737 KFFCLLTNCIBKIGQSSMW-----KALHALRPOOL----- 767
QY 757 CFFLGALTPMAVRTFAITSLGAVLDFLLQMSAPVALLSDSKRQEARLDVCCCKVPQ 816
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 768 -----TTEVLSGQAIFONILTEG-----DICASQRA 795
QY 817 ELPPPGOGEGLLGFFQKAYAPFLFHLWITRGVLLFLFALFVSLYSMCHISVGLDQELA 876
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 796 AC---EGLGLLARLGNDFIARTMI-WALHG-LLLTTEAAGLSFVSHVQAALGLALDIL 848
QY 877 LPKOSYLDYFLFLNRYFEVGAPVYVVTILGYNFSSEAGMNAICSSAGCNRNFSFTQKIY 936
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 849 LTESGWDLSQIGIRLINA-----IVAVLGPGLSPGSLFSCRKSV-IABEISWQEIPT 902
QY 937 ATERPEQSYLAIPASSWDDFDLWTPSSCCRLYISGPNKDK-----FCPSTVNS 986
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 903 LLEYVNAFLPATIILHCHYYMYNLYPVSSL-IFIVNLNIFRSVCFTQQLILFAPOAVSV 961
QY 987 LNLCLNCSITMGSRVPSV-----EQFKYLPWFIND----- 1018
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 962 HHVKNLL-MTLASQPIIRLSVSTRHLVKEKDPVSVIDEQIENLFQMLDEETDSBIG 1020
QY 1019 -----RPN-----IKCPKGLLAYS-----TSVNLTSGQ 1043
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 1021 NLIRSTLRLLYATCPSRPSRWMLICRNMAAASAGSAETSIAENDPAYTRENLGDDDE 1080
QY 1044 VLASRFMAVHKPKNSQDYTEALR-AARELAANITADLRKVPGTDP-A-FEV----- 1092
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 1081 DMVS---SSGKSIIRANPKDKTLRTRTVFAAECLSLPEAVGNDAAHFDILLARNLASN 1138
QY 1093 -----PPYITNVFYEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDL 1135
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 1139 RQSSGDLVLQLQELISLAYQISTI---QFENMRPIGVGLLSTILEKAQLLSAVRTALDA 1195
QY 1136 RSG 1138
   ||
Db 1196 NSG 1198

RESULT 86
AF3386
protein translocase chain secD / protein translocase chain secE [imported] - Brucella me
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AF3386
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Muijer, C.; Loeb, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hegius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitene
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3386
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-775 <KUR>
A:Cross-references: UNIPROT:Q8YGT1; UNIPARC:UPI0000057FID; GB:AE008917; PIDN:AAL52257.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11076
A:Map position: 1

Query Match 1.7%; Score 119.5; DB 2; Length 775;
Best Local Similarity 21.0%; Pred. No. 3.3;
Matches 120; Conservative 88; Mismatches 223; Indels 141; Gaps 27;
```

```
QY 334 HTLQGPFQCGWT---WVASWPLTILVLSVIPVVALAAGLVFTLTTPDVELWAPNSQA 390
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 16 HSMU---IFSRSKKSALIWA-----VLVSLII-----ASPNFLS 46
QY 391 RSEKAFDQHFQPFRTNQVILTAPNRSSRYRSDLLGPKNFSGILDDLLELLELQER 450
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 47 RETLA---NLPDFLPKKQVSLGLDLSGSR---LILQVQN-AGKTDLETTANI--MRQR 96
QY 451 LRHLQWNSPEAQ---RNIQLQDICYAPLNPDTNLSYDCCINSLLQYQNNRFTLLLTANQ 507
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 97 LEELGYGNPVVGEGRN-----QIRVEVPGLYDA-----QLLKDIILTIRGNL 138
QY 508 TLMQQTQSDWKD-----HFLYCANAP-----LTFKDGTTALALSCLMADY 546
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 139 SFRAMDDTMSPDPAIRGTPLADSEIVYSFDDPPVGYLLKKTPILTGHDITDAKASISADD 198
QY 547 GAPVFPFL-----AIGYKGKDYSEAEALIMFTSLNNYPAGDPRLAQAKLM-EEAF-- 596
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 199 GQPVITLTDNGRRRLADLTAQGNENSAIVVDNQVVSAPTSGPLDTSBLQIEGAFDL 258
QY 597 --LEEMRQFQRMAGMQVTFPTAERSLEDEINRTTAAEDLPFATSYIVIFLYISALGSY 654
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 259 QAANNMAVLRSGALPOAVTVLEERTIASAL---GED---YASAAVLAALAAALVWGLF 311
QY 655 SSWSRVMVSKATLGLGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVG-ADN 713
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 312 MVLSYGLGVTLVALVNNIILTAVLSLIG-----ASISLASTAGLVLTIGLAVD 362
QY 714 IFIFVLEYQRLPRRP---EPREVHIGRALGRVAPSMLLCSLEAICFFIGALTMPMAV 769
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 363 AHILIYERVRDRKGYSVQWESGFYRALSTVDANLTTLIAALVFLGSGT----V 418
QY 770 RTPALTSLGAVLDFLQMSAFVALLSDSKRQEARLDVCCCKPQELPPGQEGILL 829
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 419 HGPALTVAIGIGTTLFTLT--FTRELLTAQWVR-----TAKPKVEP-----KRRLK 462
QY 830 GFFQKAVAPFL-LHWITRGVLL---LFLALF 857
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 463 LVPTVTHPFMRLOFVLGISVLAIVVALF 494

RESULT 87
S76433
cation efflux system protein czcA-2 - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sl10142
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76433
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76433
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1075 <KAN>
A:Cross-references: UNIPROT:P74461; UNIPARC:UPI00000D3597; EMBL:D90915; GB:AB001339; NID
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: czcA-2
C:Superfamily: acriflavin resistance protein

Query Match 1.7%; Score 119.5; DB 2; Length 1075;
Best Local Similarity 20.2%; Pred. No. 5.1;
Matches 239; Conservative 140; Mismatches 417; Indels 385; Gaps 58;

QY 353 LTILVLSVIPVVALAAGL-----VFTELTDPVE-LWSAPN--- 387
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 26 LAIIVLGVFAVFSLPVDLLPSITYPRIGVRLDAPGVSPEVAVDITRPLEAALSATEGV 85
```

[illegible][illegible]


```
QY 1172 ----VSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVL 1227
Db 961 IEATLDVAVMRRLPILMTSLAIFLGMP-----LVISTG-----AGSGAQNAGVTGVM 1008
QY 1228 -GLAKAQLIQIFF 1239
Db 1009 GGMVTVTLAIF 1021

RESULT 95
C90693
acridine efflux pump [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: C90693
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <HAY>
A:CROSS-references: UNIPROT:Q8XD55; UNIPARC:UPI00000D03CA; GB:BA000007; PIDN:BA033938.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0515
A:Superfamily: acriflavin resistance protein

Query Match 1.7%; Score 117.5; DB 2; Length 1049;
Best Local Similarity 17.7%; Pred. No. 6.8;
Matches 140; Conservative 117; Mismatches 249; Indels 287; Gaps 33;

QY 624 EINRTTADLPFAFSYIVIFLYISALGYSYSSWRVMDSKATL--GLGGVAVVLG--A 679
Db 339 EVVKTLVE-----AIIIVFLVMTLFLQNF-----RATLIPTIAVPVLIGTFA 381

QY 680 VMAAMGFFSYLGRSLVILQVVPFLVLSVG--ADNIFIFVLEYQRLPRRPGPREVHIG 737
Db 382 VLAAGFP-----SINTLTFPG-WVLAIGLLVDDAIVVENVERVMAEGLPPKEATR 432

QY 738 RAIGR-----VAPSMLLCSLSBAICFFLGALFPMFAVTRTALTSLGLAVILDFLQMSAPV 792
Db 433 KSMGQIQOAGLVGIAMVLSAVFVPMFAFFGGSTGAI--YRQFSITIVSAMLSVLVALILTP 490

QY 793 ALLSDSKRQERASRLDVCCVQQLPPPGQEGILLGFFQKAYAPFLHWI-----844
Db 491 ALCA-----TMLKPIAKGDHGEKGGKGFNGFNRMFKSTHHTYDTSVGGILR 536

QY 845 TRGVLLLLFALFGVSLY-----SMCHISVGLDOE-----874
Db 537 STGRVILVLLIIVGMAYLVFVRLPSFLPDEDOGVMTVMVQLPAGATQRTOKLVNETH 596

QY 875 --LALPKDS-----YLLDYFLNLR-----892
Db 597 YLTKKNNVESVFAVNGFGFAGRGQNTGIAFVSLKMDADRPGEEKVKEAITMRATRAFS 656

QY 893 -----YFEVGAPYVFTTLYGNFS--SEAGM-----NAICSSAGCWNFSFT--931
Db 657 QIKDAMVAFNLPFAIVELGTATGDFDELIDQAGLGEKLTQARNQLLAEEAAKHPDMLTSV 716

QY 932 -----QKIQATEPEQSYLAIPAS-----SWVDDFDLWLTSSCRLVI 971
Db 717 RNLGLEDTPQFKIDIDQEAQALGVSINDINTLGAAGGGSYVNDIFID--RGRVKKVV 773

QY 972 SGNPKDKFCFTSVNSLNCNKMSITMGSVRPSVEQFHKYLFPWFLNDRPNIKCPKGLAA 1031
Db 774 MSEAKYRMLPDDIGD-----WYVRAADQGVMP---FSA 803

QY 1032 YST-----SVNLTSDGQVLASRFMAHYHKPLKNSQDYETALRAARELANITADLRKVP 1085
Db 804 FSSSRWEYSGPRLYRNGLPMSMEILQQAAPGKST---GEAMELMQLASKLPTGV-----855
```

```
QY 1086 TDPAFEVPPYITITNVFYEQYL-----TILPEGLFMLSICL-----VPT 1123
Db 856 -----GYDWTGMSYQERLSGNQAPSLYAIASLIWVFLCLAAALYESWSIPPSVMLVWEL 907
QY 1124 PAVSCLILGLDLRSGLNLLSIVMLIVDTVGFPMALWDISYNAVSLINL-----1171
Db 908 GVIALLAA--TFRGLTNDVYFQVGLLTTIGLSA-----KNAILIVEFAKDLMDKEGKGL 960
QY 1172 ----VSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVL 1227
Db 961 IEATLDVAVMRRLPILMTSLAIFLGMP-----LVISTG-----AGSGAQNAGVTGVM 1008
QY 1228 -GLAKAQLIQIFF 1239
Db 1009 GGMVTVTLAIF 1021

RESULT 96
B36938
acriflavin resistance protein acrB - Escherichia coli (strain K-12)
N:Alternate names: probable transmembrane protein acrE
C:Species: Escherichia coli
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: B36938; E64776
R:Ma, D.; Cook, D.N.; Alberti, M.; Pon, N.G.; Nikaido, H.; Hearst, J.E.
J. Bacteriol. 175, 6299-6313, 1993
A:Title: Molecular cloning and characterization of acrA and acrE genes of Escherichia c
A:Reference number: A36938; MUID:94012493; PMID:8407802
A:Accession: B36938
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <MAA>
A:CROSS-references: UNIPROT:P31224; UNIPARC:UPI0000125397; GB:U00734; NID:g392830; PIDN:
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E64776
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1049 <BLAT>
A:CROSS-references: UNIPARC:UPI0000125397; GB:AE000152; GB:U00096; NID:gi786660; PIDN:A
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: acrB; acrE
C:Function:
A:Description: multidrug efflux pump
A:Note: active form contains proteins acrA and acrB
C:Superfamily: acriflavin resistance protein
C:Keywords: transmembrane protein; transport protein
F:12-28/Domain: transmembrane #status predicted <TM1>
F:343-359/Domain: transmembrane #status predicted <TM2>
F:370-386/Domain: transmembrane #status predicted <TM3>
F:397-413/Domain: transmembrane #status predicted <TM4>
F:442-458/Domain: transmembrane #status predicted <TM5>
F:472-488/Domain: transmembrane #status predicted <TM6>
F:541-557/Domain: transmembrane #status predicted <TM7>
F:875-891/Domain: transmembrane #status predicted <TM8>
F:899-915/Domain: transmembrane #status predicted <TM9>
F:975-991/Domain: transmembrane #status predicted <TM10>
F:1012-1028/Domain: transmembrane #status predicted <TM11>

Query Match 1.7%; Score 117.5; DB 2; Length 1049;
Best Local Similarity 17.7%; Pred. No. 6.8;
Matches 140; Conservative 117; Mismatches 249; Indels 287; Gaps 33;

QY 624 EINRTTADLPFAFSYIVIFLYISALGYSYSSWRVMDSKATL--GLGGVAVVLG--A 679
Db 339 EVVKTLVE-----AIIIVFLVMTLFLQNF-----RATLIPTIAVPVLIGTFA 381

QY 680 VMAAMGFFSYLGRSLVILQVVPFLVLSVG--ADNIFIFVLEYQRLPRRPGPREVHIG 737
```

```
Db 382 VLAAGF-----SINTLTFG-WLAIGLLVDVAIVVENVERVMAEGLPKEATR 432
Qy 738 RALGR-----VAPSMLLCSLSBAICFFLIGALTPMPAVRTFALTSLGLAVILDFLLQMSAFV 792
Db 433 KSMGQTQGVGALVIAVMVSAVFPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALILTP 490
Qy 793 ALLSLDSKQAEASRLDVCCVCKPQELPPPGQEGILLGFQKAYAPFLLHWI-----844
Db 491 ALCA-----TMLAPIAKGDHGEGKGFPGFNFNMFKEKSTHYHTDVSGGILR 536
Qy 845 TRGVLLLFALFGVSLY-----SMCHTSVGLDQE-----874
Db 537 STGRVLVLIIVGWGAYLVFLRPSSFLPDEQGVFMTWVQLPAGATQRTQKVLNEVTH 596
Qy 875 -LALPKDS-----YLLDPLFLNR-----892
Db 597 YLLTKENRVESVFAVVGFGFAGRGQNTGIAFVSLKMDADRGEENKVEAITMRAFRAS 656
Qy 893 -----YFEVGAPYFVTTLGYNFS--SEAGM-----NAICSSAGCNNSFT--931
Db 657 QIKDAMVAFNLPALVELGTATGPFDELIDQAGLHEKLTQARNQLLBAAKHPDMLTSV 716
Qy 932 -----QKIQTATEPEOSYLAIPAS-----SWDDFDIDMLTPSSCCRLYI 971
Db 717 RPNGLDTPQKIDIDQEKAAQALGVSINDINTLGAAGGVSVDNFDID--RGRVKYVY 773
Qy 972 SGPNDKCKPSTVNSLNCCKMCKSITMGSVRPSVEQFHKYLWFLNDRPNICKPGGLAA 1031
Db 774 MSEAKYRMLPDIDG-----WYRAADGQWVP---FSA 803
Qy 1032 YST-----SVNLTSDBGVLSRMAXHKPLKNSQDYTEALRAARELAANITADLRKYPG 1085
Db 804 FSSRWEGVSGPLRYNGLPSEILQQAAPGKST--GEAMELQELASKLPTGV-----855
Qy 1086 TDAFEVFPYTTINVFYEQYL-----TILPEGLFMLSCL-----VPT 1123
Db 856 -----GYDWTGMSYQERLSGNQAPSLAIALSVLVLCLAALESWSIPFSVMLVVP 907
Qy 1124 FAVSCLLGLDLRGLLNLISVMTLVDTVGFMAWDISYNAVSLINI-----1171
Db 908 GVIGALLAA--TFRGLTNDVFPVQGLLTIGLSA-----KNAILIVEFAKOLMDKEGKL 960
Qy 1172 ----VSAGMSVEFVSHITRSPAISTKPTWLERAKEATISGSAVFAGVATNLPGLVL 1227
Db 961 IEATLDVAVRMLRPILMTSLAFLGVMP-----LVISTG-----AGSGAQNAGVTGM 1008
Qy 1228 -GLAKAQLIQIFF 1239
Db 1009 GGMVATVLAIPF 1021
```

RESULT 97

S23756
C;R protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 05-Mar-1994 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S23756
R;Tucker, S.J.; Tannahill, D.; Higgins, C.F.
Hum. Mol. Genet. 1, 77-82, 1992
A;Title: Identification and developmental expression of the Xenopus laevis cystic fibrosis
A;Reference number: S23756; MUID:93244789; PMID:1284470
A;Accession: S23756
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1485 <TUC>
C;Cross-references: UNIPROT:P26363; UNIPARC:UPI0000127558; EMBL:X65256; NID:964622; PIDN
C;Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding cassette
C;Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F;442-623/Domain: ATP-binding cassette homology <ABC1>
F;459-466/Region: nucleotide-binding motif A (P-loop)
F;1230-1422/Domain: ATP-binding cassette homology <ABC2>
F;1247-1254/Region: nucleotide-binding motif A (P-loop)

Query Match 1.7%; Score 117.5; DB 2; Length 1485;
Beet Local Similarity 17.3%; Pred. No. 11;
Matches 183; Conservative 156; Mismatches 375; Indels 343; Gaps 46;

```
Qy 265 AIARPOALDSFYLGQMGSLVLI-----ILCSVFAVVTILLVGFVAPADKSKM 316
Db 558 SLARAVYKADLYLLDSFSDYLDLFTKEIFEKIFESCVCKMANKTRILVTSKVEQLKRAKV 617
Qy 317 VDPKKG-----TSLSD-KLSFSTHTLLQFFQGWGTWVASWPLTILVLSVIPVALA 367
Db 618 LILHEGSCYFYGTSELEDQRPFSHLIGPDHEN-----AER 655
Qy 368 AGLVPE-----LTTDPVELWSAPNSQARSEKAFHQHFGFPFFTNQVILTAPNRSSVR 421
Db 656 RNSIITEFLRCSIDSDP-----SAVRNEVKK-KSF--KQVADFTKRSKSIINPKSRK 708
Qy 422 YDSLGLGPKNFGILDLDLLELLELQERLRLHLOWSPQAQNSISLQDICVAPLNPDNTS 481
Db 709 FSLMQKSPQMSGIEEDMPAEQGE-----RKLISVPESEQGEASLPRSNFLNTGPT---760
Qy 482 LYDCINSLLQYFQNNRTLLLLTANQTLMGQTSQVMDKHFLYCANAPLTPKQGTALALS 541
Db 761 -----FQRRRQSVL---NLMTRTS-----ISQGS---N 783
Qy 542 CMADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAALWE-----593
Db 784 AFATRNASVRK-MSVNSYSNS-----SFDLDIY---NRRLSODSILEVSEINEE 829
Qy 594 ---BAFLEEMAFQRMAGMFQVTF-TAERLEDEINRTTAEDLPFATSPVIVPLYSIL 649
Db 830 DLKECFLDDTDSQPTTWTNYTLFLTAHKNF-----IFILVCLVIFVFEV 876
Qy 650 ALGYSYSSW---SRMVDKATLGLGVAVVLGAVMAAMG---FFSYLGRSSLSVILQV 701
Db 877 AASSAWLWIIKRNAPAINMTSNENVSEVSDTSLVIVTHTSFYVYFVYVGVADSLALGI 936
Qy 702 ---VPFLVLSVGADNIFIVLEYQRLPRRPEPREVHIGRALGRVAPS-----MLL 749
Db 937 PRGLPLVHSLISVSKVLHKMLHALHAPMSTFNTMRAGRLNRPFKDTAILDILPLSI 996
Qy 750 CSLSEACIFFLALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLDSKQASRLDV 809
Db 997 PDLTQLVLIVIGAITVSLLEPYIFLATVPVIVAFILLRSYFL-----HTSQQLKQL- 1049
Qy 810 CCCVKPQELPPPGQEGILLGFQKAYAPFLLHWI--RGVLLLLFLALFGVSLY--SMC 865
Db 1050 -----SKARSPITPAHLITSLKG---LWTLRAFGRQPYFETLP 1083
Qy 866 HISVGLQELALPKDSYLLDYFLFN--RYPEVGAPVVFVTTLGVNPSSEAGMAICSSA 923
Db 1084 HKALNL-----HTANWFLYLSLRFQWTFIEMFVI-----1114
Qy 924 GCNNFSPQKIQTAYATEPPEOSYLAIPASSWVDDFDIMLTPSSCCRLYISGNKDK-----978
Db 1115 -----FFIAV-----SFISIAT-----SGAGEKVGIVL 1138
Qy 979 -FCPSTVNSLNCCKMCKSITMGSVRPSVEQFHKYLWFLNDRPNICKPGGLAAYSTSVN 1037
Db 1139 TLAMNIMNTLQWAVNA-SIDVDSLMSRSVSRIFRI-----DLP-----VEELIN 1181
Qy 1038 LTSQGVQLASRFMAYHKPLKNSQDYTEALR-----AARELAANITADLRKVPGTDA 1089
Db 1182 ENKKEEQLESEVLIY-----ENDYVKTQWVPSSGQWTKVLSANY-----IDGGNTV 1229
Qy 1090 FEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLNLISVIMIL 1149
Db 1230 LE-----NIFSLSLPQQRVG--LLG-RTSGKSKTLLSAFURL 1263
Qy 1150 VDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPT--TWLERAKEATIS 1207
Db 1264 LSTQGDIDIGVSWQTIPLQKWRKAFGVIPOKVFIFSGSIRKNLDIPYKGV---SDEELLK 1320
```


QY 1208 MGSAYPAGVAMTNLPQIL-----VLGLAKAQLI 1235
DB 1321 VTEEVGLKLIIDQFFQQLDVLDDGCVLGHGKQLV 1357

RESULT 98

B98137
hypothetical 46.1K protein in plac 3'region [imported] - Agrobacterium tumefaciens (stra
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98137
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B98137
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-631 <KUR>
A:Cross-references: UNIPROT:Q8U610; UNIPARC:UPI00000D2701; GB:AE007870; PIDN:AAK88620.1;
C:Genetics:
A:Gene: AGR_L113
A:Map position: linear chromosome

Query Match 1.7%; Score 117; DB 2; Length 631;
Best Local Similarity 20.4%; Pred. No. 3.7;
Matches 92; Conservative 65; Mismatches 175; Indels 120; Gaps 17;

QY 500 LLLLTANTQIMQTSQVDKDHFLYCANAPLTFKDGTA-LALSCMADYGAPV-----550

DB 46 LLAMMIGLLLVGTSR-----YALHLPYIWDIETASLCLFWFAMLGAAIAIDRSEHL 97

QY 551 -----PP-----PLAIGYKGYSEAEALIMTSLN-----577

DB 98 RLTLFLNFPQVILGYNSLALVATFLAALIKPAMEYAEIWEVWTSALNIPMSFRAA 157

QY 578 NYPAGD---PRLAQAKLMEAELEEMRAFQRMAGMFQVTFTAERSLEDEINRTAEDLP 634

DB 158 ALPVGACLLMLLVNLPRRNLRIIAAFVTVVAAGLLYLASPMLESGLNLALIFLG 217

QY 635 IPATSYIVFLYISALG-----SYSSW-SRWVWDSKATIGLGG-----VAVVIGA 679

DB 218 LFVAVFLVGVPIAFCFGTLGTLAYLTFTTWTPTVIMIGRMDGMSGILLSVFVFLGC 277

QY 680 VMAAMG-----PFSYIGIRSSVLIVQVPLVLSVG-----ADNIFIFVLEYQRL 724

DB 278 VLDATGMGKAIVELLSMPFHIRAGMSYVLLGSL-FLVSGISGSKVSDMATVAPALFPEM 336

QY 725 PRPQGPREVHIGRALG-----RVAFSMLLCSLSEALCFPLGALTTPMPAVRTFALTSGLA 779

DB 337 KRGHKPKEMIALLATGAAMADTVPPSIVLIVLSAAGVSIAGL-----FTSGPM 386

QY 780 VILDFLLQMSAFVALLSDSKQERASRLDCCCVKQFQELPPPGQEGLLLGFFQKAYAPF 839

DB 387 IAMVLLLVLAFLARWAKARDENMEGAKR-----TPWQVQVGAALIA-APALVLPF 434

QY 840 LLHWITRG-----VLLLLPLALFGVSLY 862

DB 435 LIRSLVGGGVATATEVSTIAVLYAMIVGAVLY 466

RESULT 99

T38495
hypothetical protein SPAC29B12.07 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38495
R:Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21797
A:Accession: T38495
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-1969 <GEN>
A:Cross-references: UNIPROT:O14029; UNIPARC:UPI000017B1B4; EMBL:Z99164; PIDN:CAB16252.1
A:Experimental source: strain 972h-; cosmid c29B12
C:Genetics:
A:Gene: SPDB:SPAC29B12.07
A:Map position: 1
A:Introns: 664/2

Query Match 1.7%; Score 117; DB 2; Length 1969;
Best Local Similarity 18.1%; Pred. No. 17;
Matches 204; Conservative 117; Mismatches 390; Indels 418; Gaps 46;

QY 92 SAKQJVSLEASISITKALLTRCPACSDNFVNLHCHTCSNQSLSFINVTRVLAQLGAGQAP 151

DB 806 SAMSPGNLQRTANLYKPMTT---PNAYNIKNSQRSTKYYPQPAINYSEVTPQSSSLP 862

QY 152 AVVAYEAFYQHSFAEQSYDSCSRVRVPAATLAVGTMCGVGSALCNAQRLN-----204

DB 863 T-----SGEANIIRSPGFTPLAAQKDATIYTPSHAQATLYGNMNDNRD 907

QY 205 -----FQDGTGNGLAPLDITFHLLEPGQAVGSG-----IQPLNEGVARC---NESQG 248

DB 908 NEGIIHDIQSDMEPVLPHNSAYHANAPVSSHSEGLNNELPISPLPQLHKTGTSHQHG 967

QY 249 DDVATSCQDCAASCAPAIARPOALDSTFYLGMQPSGLVLIILLCVFAVVTILLVGRVA 308

DB 968 PDAETATKQVAPSIPPNPNVSDITWTEGVALPSATL-----1006

QY 309 PARDKSKWVDPKKGTSL-----DKLSFSTHTLLGQPFQGWGTWASWPLTILVLSVI 361

DB 1007 ---DSKSSLHKRSALSRNNSPRDPFLPLNQPLLH-----SNLHSPVS 1048

QY 362 PVVLAAGLVFTLTTPDVELWSAPNSOAR-----SEKAFHQHFGPPFTNQVILTAP 415

DB 1049 PV-----DSEDSRLKFLSTQPAF---SFGPC---GTIVMAF 1081

QY 416 NRSSRYDLSLLGPKNFSGILDLLLELLELQERLHLQVMSPEAQRNLSLQDICYAPL 475

DB 1082 STPSGLYTTSGKGTGFIAGPIKIEKLDV---LTDEYRHLKBFKG-----1123

QY 476 NPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKD- 534

DB 1124 -----PYLASN-----GKVDKHKAEAEIEMLSKYIDRLAQSLSEYDDK 1160

QY 535 -----GTALALSCMADYGAPVFPFLAIGYKGYSEAEALIMTSLNNYPAGDPRLAQA 589

DB 1161 NITLKDQLLLQCL-----KMLEVS-----DRKL-----1185

QY 590 KLWEEAFLEEMRAFQRMAGMFQV-----TFTAERSL-EDEINRTTAEDLPFATSYIVI 643

DB 1186 -----IVKELRPI---LLPSPEIPEPCNTATSVQELINPEINQ---DDSPIVASRYCTT 1233

QY 644 -FLYISLALGYSYSSRWVWDSKATIGLGVAVLVGAVMAAMGPFYSYLGIIRSSVLQV 702

DB 1234 SPLH-----RFEYL-----1243

QY 703 PELVLSVGADNIFIFVLEYQRLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFPLCA 762

DB 1244 ---LSGNKDEALTYALQOKOMP-----YALIVAHSIDA 1273

QY 763 LTPMPAVRTFA-----LTSGLAVILDFLQMSAFVALLSDSKQERASRLDCCCVK 814

DB 1274 KTFQGVRTFCKSEVKESMLRSGVGNLQSLQMSDAHASSMSFSSSTSLN-----1327

QY 815 POELPPPGQEGLLGFFQKAYAPFLHWITR-----GVVLLPLALFGVLSVSMCHI 867

DB 1328 ---LADQASNALVANKELLYNIINAHYSDQKEALVGLTLLQENRVAAHLVYVLSL 1384

QY 868 S-----VGLDOELALPKDSVILDDYF-----LFNRYPEVGVAPVFTVL 906

DB 1385 SPDVCNKSNSLFLVGLSKHNLPSHDDLDPVLTQTEVLVFNVFNYS-KTFVFPFTHLV 1443

Qy 907 GYNE-----SSEAG-----MNAICSSAGCNFSFTQKIQVATEPEEQSYL 946
Db 1444 PYRIEAEVLAEGEVSARKYCELIYNLRVAKSNVNDPGFVLVRDLT-----QQIL 1499
Qy 947 AIPA-----SSWVD-----DFIDWLTSSCCRLYISGPNKDKFCPSTV-----NSL 987
Db 1500 ENSAGSEDISSWLGRTVSRPLDVLSSLGSKFSKFVAGDPNFDVVRPATVGPFGK 1559
Qy 998 NCLKNCMSITGWSVRPSVEQPHKYLFWFLNDRPNIKCPKGGLAAYSTSVNLTSDGOVLAS 1047
Db 1560 ASQKN-LTVQNTNNAAMESPY-----SDRPTSSGP-----SYQNRTPLT--GQ--ES 1602
Qy 1048 RFMAVHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYT 1096
Db 1603 MMGVSPYRRS-----TEIAENMSDGNAYPYT-PASQENPYT 1640

RESULT 100
AE0380
multidrug efflux protein [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE0380
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0380
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1050 <KUR>
A:Cross-References: UNIPROT:Q8ZC87; UNIPARC:UPI00000DCB3F; GB:AL590842; PIDN:CAC92368.1;
C:Genetics:
A:Gene: acrB
C:Superfamily: acriflavin resistance protein

Query Match 1.7%; Score 116; DB 2; Length 1050;
Best Local Similarity 17.7%; Pred. No. 8.7; Mismatches 243; Indels 310; Gaps 35;
Matches 144; Conservative 115;

Qy 623 DEINKTARDLPFATSYIVIEFLYTSIALGYSSSRVMSVDSKATL-GLGGVAVVLGAVM 681
Db 338 NEVKTLE-----AIIIVLVMVLFQNF-----RATLITPIAVPVLLGTF 380
Qy 682 AAMGFYSYLGIRSLVILQVPFLVLSVG--ADNIFIVLEYQRLPRRPGEPREVIHGRA 739
Db 381 AILSAFGY-----SINTLTMPG-MVLAIGLLVDDAIVVVENVERVMQBEGLPPKEATKKS 434
Qy 740 LGR-----VAPSMLLCSLSEATCFPLGALTMPAVRTPALTSLGLAVILDFLLQMSAFVAL 794
Db 435 MEQIQGALVIALVLSAVFVPAFPGATGAI--YRQFSITIVSAMVLSVALILITPAL 492
Qy 795 LSLSKRSQASRLDVCCKVQBELPPQGGELLLGFFQKAYAPFLHMI-----TR 846
Db 493 CA-----TMLKPIKGDHGPKTG-FFGWFNNMPEKSTHHYTDSDVANILRST 537
Qy 847 GVVLFLFLALF-----GVSLYSMCHI SVGLDQELALPKDSYLLDY 886
Db 538 GRYLVIYLAIVIGMAVLFMLRPFSSFLPEBDQGVFL-TMVQLPAGATQERTQKVLNHVTDY 596
Qy 887 FL-----FLNRYFEVCAPYVFTTLGYNESSEAGMN-----AI 919
Db 597 YLDKEKNVNVSVPTN-----GFGFSQG-QGNTGLAFVSLKNWDERKGEQNKVPAL 646
Qy 920 CSSAG-----CWNFSF-----TKIQYAT 938
Db 647 VSEASAFSKIKDGMVFAFNLPALPAVELGTATGDFQLIDQGNLGHQQLTDARNQLLGMAA 706
Qy 939 EFPEQSYLAIP-----ASSWDDDFIDWLT 962
Db 707 QHPDMLVGRPNGLGLEDTPQKVEVDQEKQAQALGVAISDINTTILGSAAGGSYVNDFFD---- 763

Qy 963 PSSCCRLYISGPNKDKFCPSTVNSLNCNCKNSITMGSVRPSVEQPHKYLFWFLNDRPNI 1022
Db 764 RGRVKVYVQADAPFRMLPDDIDK-----WYVRNN--- 793
Qy 1023 KCPKGGLAAYST-----SVNLTSDGOVLASRFMAYHKPLKNSODYTEALRAARELA 1073
Db 794 ---MGQMVSPATFSTAKWEYGSPLRYNGLPSPMEILGQAAPGKST---GEAMDLMQELA 847
Qy 1074 ANITADLRKVPCTDPAFEVFPYTITNVFYEQYLT-----ILPEGLFMLSICL----- 1120
Db 848 AKLP SGV-----GYDWTGMSYQERLSGNQAPADYAISSLIVVFLCLLAALYESW 894
Qy 1121 -----VPTFAVSCLLILGLDLRLSGLLMLLSITVILVDTVGFPMALWDISYNAVSLI-- 1169
Db 895 SIPPSVMLVPLGVVGA-LLAATLR-GLENDVYFQVGLLTITIGLSA-----KNAILIIVEF 947
Qy 1170 -----NLVSAGVMSVEFVSHITRSPTAISTKPTWLERAKEATISMGSAPVAG 1215
Db 948 AKDLMDKEGKGLVESTLESVRMLRPIILMTSLAFILGVMP-----LVISSG-----AG 995
Qy 1216 VAMTNLPGLVL-GLAKAQLIQIFFFRLNLLI 1246
Db 996 SGAQNAVGTGVMGMITATVLAIFVFLFFV 1027

Search completed: April 11, 2006, 00:46:05
Job time : 92 secs

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: April 11, 2006, 00:36:46 ; Search time 265 Seconds
 (without alignments)
 3546.278 Million cell updates/sec

Title: US-10-736-769-4
 Perfect score: 6909
 Sequence: 1 MAAEGLRWLLWALLRLAQ.....GSIKAGALSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6872.5	99.5	1359	1 NPCL1_HUMAN	Q9ubc9 homo sapien
2	5421.5	78.5	1331	1 NPCL1_RAT	Q6t3u3 rattus norv
3	5407	78.3	1333	1 NPCL1_MOUSE	Q6t3u4 mus musculu
4	3179	46.0	1132	2 Q4T749_TETNG	Q4t749 tetraodon n
5	2413	34.9	1277	2 Q9JLG3_CRIGR	Q9jlg3 cricetus
6	2412	34.9	1277	1 NPCL1_PTG	P56941 sus scrofa
7	2405	34.8	1276	2 Q8MKD8_FELCA	Q8mkd8 felis silve
8	2404.5	34.8	1276	2 Q9GK52_CANFA	Q9gk52 canis fami
9	2402.5	34.8	1278	1 NPCL1_HUMAN	O15118 homo sapien
10	2402	34.8	1289	2 Q59GR1_HUMAN	Q59gr1 homo sapien
11	2400	34.7	1277	2 Q7TMD4_MOUSE	Q7tmd4 mus musculu
12	2399	34.7	1276	2 Q9NQO0_FELCA	Q9nqo0 felis silve
13	2395	34.7	1276	2 Q8MI49_FELCA	Q8mi49 felis silve
14	2385	34.5	1278	1 NPCL1_MOUSE	Q35604 mus musculu
15	2367.5	34.3	1277	2 Q9GLC9_BOVIN	Q9glc9 bos taurus
16	2361	34.2	1286	2 Q9TT75_RABIT	Q9tt75 carytolagus n
17	2338	33.8	1209	2 Q4RWY5_TETNG	Q4rwy5 tetraodon n
18	2262	32.7	1287	2 Q9VL24_DROME	Q9vl24 drosophila
19	2256	32.7	1287	2 Q9U5W1_DROME	Q9u5w1 drosophila
20	2253	32.6	1287	2 Q7YU59_DROME	Q7yu59 drosophila
21	2205	31.9	1291	2 Q7YU09_ANOGA	Q7yu09 anopheles g
22	1864.5	27.0	1233	2 Q9VRC9_DROME	Q9vrc9 drosophila
23	1760	25.5	1003	2 Q7PSO3_ANOGA	Q7ps03 anopheles g
24	1725	25.0	1275	2 Q5LNK7_MAGGR	Q5lnk7 magnaporth
25	1712.5	24.8	1361	2 Q7XUB7_ORYSA	Q7xub7 oryza sativ
26	1700.5	24.6	1271	2 Q5BBG1_EMENI	Q5bbg1 aspergillus
27	1659.5	24.0	1330	2 Q5KGS9_CRYNE	Q5kgs9 cryptococcu
28	1658.5	24.0	1330	2 Q5SSD4_CRYNE	Q5ssd4 cryptococcu
29	1643	23.8	1273	2 Q4WNG5_ASFFU	Q4wns5 aspergillus
30	1614.5	23.4	1295	2 Q41OK4_GIBBE	Q41ok4 gibberella
31	1549	22.4	1264	2 Q6BT03_DEBHA	Q6bt03 debaryomyce

ALIGNMENTS

32	1527.5	22.1	1275	2 Q9SHN9_ARATH	Q9shn9 arabidopsis
33	1511.5	21.9	1342	2 Q9TVK6_DICDI	Q9tvk6 dictyosteli
34	1508	21.8	1239	2 Q6CBAL_YARLI	Q6cbal yarrowia li
35	1500	21.7	1162	2 Q7RWL9_NEUCR	Q7rwl9 neuropept
36	1461.5	21.2	1256	2 Q592V0_CANAL	Q592v0 candida alb
37	1452.5	21.0	1397	2 Q551C5_DICDI	Q551c5 dictyosteli
38	1437.5	20.8	1489	2 Q4PE83_USTMA	Q4pe83 ustilago ma
39	1394	20.2	1055	2 Q9SVF0_ARATH	Q9svf0 arabidopsis
40	1386.5	20.1	620	2 Q4H344_CIOIN	Q4h344 ciona intes
41	1359	19.7	1382	2 Q618V4_CAEBR	Q618v4 caenorhabdi
42	1341.5	19.4	1383	1 NPCL1_CAEBL	Q19127 caenorhabdi
43	1329	19.2	1170	2 Q12200_YEAST	Q12200 saccharomyc
44	1323	19.1	1339	2 Q5ORB6_ENTHI	Q5orb6 entamoeba h
45	1312.5	19.0	1178	2 Q750G1_ASHGO	Q750g1 ashbya goss
46	1300.5	18.8	1177	2 Q6CIF2_KLULA	Q6cif2 kluyveromyc
47	1274.5	18.4	1174	2 Q6FKT2_CANGA	Q6fkt2 candida gla
48	1046	15.1	1274	1 NPCL1_CAEBL	Q19127 caenorhabdi
49	981	14.2	1222	2 Q612M9_CAEBR	Q612m9 caenorhabdi
50	876	12.7	371	2 Q6GNS9_CIOIN	Q66nr8 ciona intes
51	874	12.7	371	2 Q6GNS8_CIOIN	Q66nr8 ciona intes
52	873	12.6	371	2 Q6GNS4_CIOIN	Q66nr4 ciona intes
53	873	12.6	371	2 Q6GNS3_CIOIN	Q66nr4 ciona intes
54	873	12.6	371	2 Q6GNS2_CIOIN	Q66nr4 ciona intes
55	869	12.6	371	2 Q6GNS1_CIOIN	Q66nr4 ciona intes
56	869	12.6	371	2 Q6GNS0_CIOIN	Q66nr4 ciona intes
57	867	12.5	371	2 Q6GNT1_CIOIN	Q66nr4 ciona intes
58	858.5	12.4	524	2 Q9CF80_ARATH	Q9cf80 arabidopsis
59	854	12.4	371	2 Q6GNS0_CIOIN	Q66nr4 ciona intes
60	776.5	11.2	601	2 Q5TQH9_ANOGA	Q5tqh9 anopheles g
61	651.5	9.4	1447	1 PTC1_HUMAN	Q13835 homo sapien
62	649	9.4	1434	2 Q6UY90_RAT	Q6uy90 rattus norv
63	646.5	9.4	1434	1 PTC1_MOUSE	Q61115 mus musculu
64	636.5	9.2	827	2 Q5RJ13_BRARE	Q5rj13 brachydanio
65	628.5	9.1	1442	1 PTC1_CHICK	Q5rj13 gallus gall
66	624	9.0	1296	2 Q5VZC2_HUMAN	Q5vzc2 homo sapien
67	617.5	8.9	1239	2 Q9DEF4_XENLA	Q9def4 xenopus lae
68	617.5	8.9	1418	2 Q98SW6_XENLA	Q98sw6 xenopus lae
69	601.5	8.7	1243	2 Q9W6T6_BRARE	Q9w6t6 brachydanio
70	599.5	8.7	1413	2 Q9DEF3_XENLA	Q9def3 xenopus lae
71	598.5	8.7	1422	2 Q6IRA5_XENLA	Q6ira5 xenopus lae
72	598.5	8.7	933	2 P91129_CAEBL	P91129 caenorhabdi
73	597	8.6	896	2 Q617L5_CAEBL	Q617l5 caenorhabdi
74	592	8.6	900	2 P91346_CAEBL	P91346 caenorhabdi
75	589	8.5	278	2 Q4H343_CIOIN	Q4h343 ciona intes
76	585	8.5	933	2 Q61EN1_CAEBR	Q61en1 caenorhabdi
77	584	8.5	1203	1 PTC2_HUMAN	Q9vsc5 homo sapien
78	584	8.5	1203	2 Q53Z57_HUMAN	Q53z57 homo sapien
79	583.5	8.4	1182	1 PTC2_MOUSE	Q35595 mus musculu
80	573.5	8.3	1003	2 Q9XXR9_CAEBL	Q9xxr9 caenorhabdi
81	572	8.3	1220	1 PTC1_BRARE	Q98864 brachydanio
82	566.5	8.2	1004	2 Q60VB2_CAEBR	Q60vb2 caenorhabdi
83	562.5	8.1	1169	2 Q9V9L4_DROME	Q9v9l4 drosophila
84	532	7.7	816	2 Q4SP00_TETNG	Q4sp00 tetraodon n
85	522.5	7.6	1318	2 Q9XYP5_JUNCO	Q9xyp5 junonia coe
86	522	7.5	1197	2 Q61232_CAEBR	Q61232 caenorhabdi
87	521.5	7.5	1405	2 Q4RVD6_TETNG	Q4rvd6 tetraodon n
88	514	7.4	1405	1 PTC1_CAEBL	Q09614 caenorhabdi
89	474	6.9	1286	1 PTC3_DROME	P18502 drosophila
90	473.5	6.9	1015	2 Q09938_CAEBL	Q09938 caenorhabdi
91	470.5	6.8	915	2 Q61PC1_CAEBR	Q61pc1 caenorhabdi
92	464	6.7	974	2 Q621T1_CAEBR	Q621t1 caenorhabdi
93	463	6.7	1380	2 Q60VE7_CAEBR	Q60ve7 caenorhabdi
94	458.5	6.6	955	2 Q19945_CAEBL	Q19945 caenorhabdi
95	455	6.6	1358	2 Q6AW15_CAEBL	Q6aw15 caenorhabdi
96	453.5	6.6	1361	2 Q6AW16_CAEBL	Q6aw16 caenorhabdi
97	444.5	6.4	844	1 PTC9_CAEBL	Q03602 caenorhabdi
98	436	6.3	956	2 Q18129_CAEBL	Q18129 caenorhabdi
99	424	6.1	951	2 Q60W90_CAEBR	Q60w90 caenorhabdi
100	422.5	6.1	939	2 Q7Q2Y4_ANOGA	Q7q2y4 anopheles g

RESULT 1
ID NPCL1 HUMAN STANDARD; PRT; 1359 AA.
AC Q9UHC9; Q6R304; Q9UHC9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Niemann-Pick C1-like protein 1 precursor.
GN Name=NPCL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 2 AND 3), AND TISSUE SPECIFICITY.
RX MEDLINE=20247253; PubMed=10783261; DOI=10.1006/geno.2000.6151;
RA Davies J.P., Levy B., Ioannou Y.A.;
RT Evidence for a Niemann-Pick C (NPC) gene family: identification and characterization of NPCL1.";
RL Genomics 65:137-145(2000).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), AND TISSUE SPECIFICITY.
RX PubMed=14976318; DOI=10.1126/science.1093131;
RA Altmann S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
RA Tetzlaff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
RA Murgolo N., Graziano M.P.;
RT "Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol absorption.";
RL Science 303:1201-1204(2004).
[3]
RN TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15671032; DOI=10.1074/jbc.M409110200;
RA Davies J.P., Scott C., Oishi K., Liapis A., Ioannou Y.A.;
RT "Inactivation of NPCL1 causes multiple lipid transport defects and protects against diet-induced hypercholesterolemia.";
RL J. Biol. Chem. 280:12710-12720(2005).
[4]
RN INDUCTION.
RX PubMed=15604518; DOI=10.1194/jlr.M400400-JLR200;
RA van der Veen J.N., Kruit J.K., Havings R., Baller J.F.W., Chimini G.,
RA Lescavel S., Staelens B., Groot P.H.E., Groen A.K., Kuipers F.;
RT "Reduced cholesterol absorption upon PPARdelta activation coincides with decreased intestinal expression of NPCL1.";
RL J. Lipid Res. 46:526-534(2005).
[5]
RN FUNCTION.
RX PubMed=15928087; DOI=10.1073/pnas.0500269102;
RA Garcia-Calvo M., Lisnock J., Bull H.G., Hawes B.E., Burnett D.A.,
RA Braun M.P., Crona J.H., Davis J.H. Jr., Dean D.C., Demers P.A.,
RA Graziano M.P., Hughes M., Macintyre D.E., Ogawa A., O'Neill K.A.,
RA Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,
RA Ujjainwalla F., Altmann S.W., Chapman K.T., Thornberry N.A.;
RT "The target of ezetimibe is Niemann-Pick C1-like 1 (NPCL1).";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137(2005).
[6]
RN VARIANTS LEU-55 AND ASN-1233.
RX PubMed=15679830; DOI=10.1111/j.1399-0004.2004.00388.x;
RA Wang J., Williams C.M., Hegele R.A.;
RT "Compound heterozygosity for two non-synonymous polymorphisms in NPCL1 in a non-responder to ezetimibe.";
RL Clin. Genet. 67:175-177(2005).
CC -1- FUNCTION: Play a major role in cholesterol homeostasis. Is critical for the uptake of cholesterol across the plasma membrane of the intestinal enterocyte. Is the direct molecular target of ezetimibe, a drug that inhibits cholesterol absorption. Lack of activity leads to multiple lipid transport defects. The protein may have a function in the transport of multiple lipids and their homeostasis, and may play a critical role in regulating lipid metabolism.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation

of brush border membranes from proximal enterocytes suggests considerable association with the apical membrane fraction. Exists as a predominantly cell surface membrane expressed protein (By similarity). According to Ref.3 localizes in a subcellular vesicular compartment rich in RAB5.
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Name=1;
IsoId=Q9UHC9-1; Sequence=Displayed;
Name=2; Synonyms=NPCL1DELTAEL5;
IsoId=Q9UHC9-2; Sequence=VSP_015314;
Name=3; Synonyms=NPCL1T;
IsoId=Q9UHC9-3; Sequence=VSP_015312, VSP_015313;
TISSUE SPECIFICITY: Widely expressed. Expressed in liver. Also expressed in small intestine, pancreas, kidney, lung, pancreas, spleen, heart, gall bladder, brain, testis, stomach and muscle.
-1- INDUCTION: Expression is decreased in Caco-2 cells upon PPARd activation.
-1- PTM: Highly glycosylated (By similarity).
-1- POLYMORPHISM: Variations in NPCL1 gene could be associated with nonresponse to ezetimibe treatment
-1- SIMILARITY: Belongs to the patched family.
-1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AF192522; AAF20396.1; -; mRNA.
EMBL; AF192523; AAP20397.1; -; mRNA.
EMBL; AY515256; AAS56939.1; -; mRNA.
EMBL; AY437865; AAR57886.1; -; mRNA.
DR Ensembl; ENSG00000015520; Homo sapiens.
HGNC; HGNC:7898; NPCL1.
MIM; 608010; -;
GO; GO:0015918; P:sterol transport; TAS.
InterPro; IPR004765; NP_C type.
InterPro; IPR003392; Patched.
InterPro; IPR000731; SSD_5TM.
Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
PROSITE; PS50156; SSD; 1.
KW Alternative splicing; Cholesterol metabolism; Glycoprotein; Lipid metabolism; Polymorphism; Signal; Steroid metabolism; Transmembrane.
FT SIGNAL 1 21 Potential.
FT CHAIN 22 1359 Niemann-Pick C1-like protein 1.
FT TOPO_DOM 22 284 Extracellular (Potential).
FT TRANSMEM 285 305 1 (Potential).
FT TOPO_DOM 306 351 Cytoplasmic (Potential).
FT TRANSMEM 352 372 2 (Potential).
FT TOPO_DOM 373 632 Extracellular (Potential).
FT TRANSMEM 633 653 3 (Potential).
FT TOPO_DOM 654 666 Cytoplasmic (Potential).
FT TRANSMEM 667 687 4 (Potential).
FT TOPO_DOM 688 696 Extracellular (Potential).
FT TRANSMEM 697 717 5 (Potential).
FT TOPO_DOM 718 742 Cytoplasmic (Potential).
FT TRANSMEM 743 763 6 (Potential).
FT TOPO_DOM 764 776 Extracellular (Potential).
FT TRANSMEM 777 797 7 (Potential).
FT TOPO_DOM 798 846 Cytoplasmic (Potential).
FT TRANSMEM 847 867 8 (Potential).
FT TOPO_DOM 868 882 Extracellular (Potential).
FT TRANSMEM 883 903 9 (Potential).
FT TOPO_DOM 904 1139 Cytoplasmic (Potential).
FT TRANSMEM 1140 1160 10 (Potential).
FT TOPO_DOM 1161 1168 Extracellular (Potential).
FT TRANSMEM 1169 1189 11 (Potential).
FT TOPO_DOM 1190 1191 Cytoplasmic (Potential).
FT TRANSMEM 1192 1212 12 (Potential).

FT	TOPO_DOM	1213	1236	Extracellular (Potential).
FT	TRANSMEM	1237	1257	13 (Potential).
FT	TOPO_DOM	1258	1359	Cytoplasmic (Potential).
FT	DOMAIN	632	797	SSD.
FT	COMBIAS	500	503	Poly-Leu.
FT	CARBOHYD	54	54	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	132	132	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	138	138	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	244	244	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	416	416	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	431	431	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	464	464	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	479	479	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	497	497	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	506	506	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	626	626	N-linked (GlcNAc. .) (Potential).
FT	VARSPLIC	723	724	RL -> GP (in isoform 3).
FT				/FTid=VSP 015312.
FT	VARSPLIC	725	1359	Missing (in isoform 3).
FT				/FTid=VSP 015313.
FT	VARSPLIC	1046	1072	Missing (in isoform 2).
FT				/FTid=VSP 015314.
FT	VARIANT	55	55	V -> L (nonresponse to ezetimibe treatment).
FT				/FTid=VAR 023369.
FT	VARIANT	1233	1233	I -> N (nonresponse to ezetimibe treatment).
FT				/FTid=VAR 023370.
SQ	SEQUENCE	1359 AA; 148698 MW; 3225D53D93B768B6 CRC64;		
Query Match				
Best Local Similarity 99.5%; Score 6872.5; DB 1; Length 1359;				
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;				
QY	1	MAEAGLRGWLWALLRLAQAQSPYTTIHOPGCAFYDECGKPELSGSLMTLSNSCLSN	60	
DB	1	MAEAGLRGWLWALLRLAQAQSPYTTIHOPGCAFYDECGKPELSGSLMTLSNSCLSN	60	
QY	61	TPARKITGDHLILLOKICPRLYTGPNTOACCSAKOLVLSLEASLSITKALLTRCPACSDNF	120	
DB	61	TPARKITGDHLILLOKICPRLYTGPNTOACCSAKOLVLSLEASLSITKALLTRCPACSDNF	120	
QY	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA	180	
DB	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA	180	
QY	181	ATLAVGTMCVYGSALCNAQRWLNFGDGTNGGLAPLDITPHLLPEQAVGSGIQPLNEG	240	
DB	181	ATLAVGTMCVYGSALCNAQRWLNFGDGTNGGLAPLDITPHLLPEQAVGSGIQPLNEG	240	
QY	241	ARCNESQGDVATSCQDCCAAPARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI	300	
DB	241	ARCNESQGDVATSCQDCCAAPARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI	300	
QY	301	LLVGFVRVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWELTILVLSV	360	
DB	301	LLVGFVRVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWELTILVLSV	360	
QY	361	IPVVALAAGLVTELTDPVELWSPAPNSQARSEKAFHDQHPFRFTNQVILTAENRSY	420	
DB	361	IPVVALAAGLVTELTDPVELWSPAPNSQARSEKAFHDQHPFRFTNQVILTAENRSY	420	
QY	421	RYDSLILGPKNFGSLDLDLLELLEQLERLHQLVMSPEAQRNLSLODICVAPLNPDNT	480	
DB	421	RYDSLILGPKNFGSLDLDLLELLEQLERLHQLVMSPEAQRNLSLODICVAPLNPDNT	480	
QY	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTKDGTALAL	540	
DB	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTKDGTALAL	540	
QY	541	SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLBEM	600	
DB	541	SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLBEM	600	

QY	601	RAFQRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIIVIFLYISIALGSSWSRV	660	
DB	601	RAFQRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIIVIFLYISIALGSSWSRV	660	
QY	661	MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVSGADNIFIVLE	720	
DB	661	MYDSKATILGCGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVSGADNIFIVLE	720	
QY	721	YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTTALTSLGLAV	780	
DB	721	YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTTALTSLGLAV	780	
QY	781	ILDPLLOMSAFVALLSLDSKQEAASRLDVCCCKPQELPPPPQGGEGLLGFFQKAYAPFL	840	
DB	781	ILDPLLOMSAFVALLSLDSKQEAASRLDVCCCKPQELPPPPQGGEGLLGFFQKAYAPFL	840	
QY	841	LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900	
DB	841	LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900	
QY	901	YFVTTTLGYNFSSEAGMNAICSSAGCNFSTOKIOVATPEPQSYLAIPASSWDDFDIOW	960	
DB	901	YFVTTTLGYNFSSEAGMNAICSSAGCNFSTOKIOVATPEPQSYLAIPASSWDDFDIOW	960	
QY	961	LTPSSCCRLYISGPNKDKFCPSTVNSLCLNCKMSITMGSVRPSVEQFHKYLFPWFLNDRP	1020	
DB	961	LTPSSCCRLYISGPNKDKFCPSTVNSLCLNCKMSITMGSVRPSVEQFHKYLFPWFLNDRP	1020	
QY	1021	NIKCPKGGIAAYSTVNLTSDGQVL-----ASRPMAYH 1053		
DB	1021	NIKCPKGGIAAYSTVNLTSDGQVL-----ASRPMAYH 1053		
QY	1054	KPKNSQDYTRALRAARELAANITADLRKVPCTDPAPFPVYTTITNVFEQVLTILPEGL	1113	
DB	1081	KPKNSQDYTRALRAARELAANITADLRKVPCTDPAPFPVYTTITNVFEQVLTILPEGL	1140	
QY	1114	PMLSICLVPTFAVSCLLGLDLRLSGLNLLSVMTILVDTVGFMAWLDISYNAVSLINLVS	1173	
DB	1141	PMLSICLVPTFAVSCLLGLDLRLSGLNLLSVMTILVDTVGFMAWLDISYNAVSLINLVS	1200	
QY	1174	AVGMSVEFVSHITRFAISTKPTWLERAKBATISMGSAVFAVAGVAMTNLPGILVLGLAKAQ	1233	
DB	1201	AVGMSVEFVSHITRFAISTKPTWLERAKBATISMGSAVFAVAGVAMTNLPGILVLGLAKAQ	1260	
QY	1234	LIQIEFFRLNLLITLLGLLHGLVFLPVILSVYGPVDPNPALEOKRAEAAVAVMVASCP	1293	
DB	1261	LIQIEFFRLNLLITLLGLLHGLVFLPVILSVYGPVDPNPALEOKRAEAAVAVMVASCP	1320	
QY	1294	NHPSRVSTADNITYNHSPEGSIKGAGAINFLPNNGRQF 1332		
DB	1321	NHPSRVSTADNITYNHSPEGSIKGAGAINFLPNNGRQF 1359		

RESULT 2

NPCL1 RAT	STANDARD;	PRT; 1331 AA.
ID NPCL1 RAT		
AC O6T3U3;		
DT 13-SEP-2005 (Rel. 48, Created)		
DT 13-SEP-2005 (Rel. 48, Last sequence update)		
DT 13-SEP-2005 (Rel. 48, Last annotation update)		
DE Niemann-Pick C1-like protein 1 precursor.		
GN Name=Npc1l1;		
OS Rattus norvegicus (Rat).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC Murioidea; Muridae; Murinae; Rattus.		
OX NCBI TaxID=10116;		
RN (1)		
RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND SUBCELLULAR		
RC LOCATION.		
RC STRAIN=Sprague-Dawley;		
RX PubMed=14976318; DOI=10.1126/science.1093131;		

RA Altman S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
 RA Tetzloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
 RA Murgolo N., Graziano M.P.;
 FT "Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol
 RT absorption.";
 RL Science 303:1201-1204(2004).
 RN [2]
 RP SUBCELLULAR LOCATION, AND CHARACTERIZATION.
 RX PubMed=15777641; DOI=10.1016/j.bagen.2004.12.021;
 RA Iyer S.P.N., Yao X., Crona J.H., Hoos L.M., Tetzloff G.,
 RA Davis H.R. Jr., Graziano M.P., Altman S.W.;
 RA "Characterization of the putative native and recombinant rat sterol
 RT transporter Niemann-Pick C1 Like 1 (NPC1L1) protein.";
 RL Biochim. Biophys. Acta 1722:282-292(2005).
 RN [3]
 RP FUNCTION.
 RX PubMed=15928087; DOI=10.1073/pnas.0500269102;
 RA Garcia-Calvo M., Lissack J., Bull H.G., Hawes B.E., Burnett D.A.,
 RA Braun M.P., Crona J.H., Davis H.R. Jr., Dean D.C., Detmers P.A.,
 RA Graziano M.P., Hughes M., MacIntyre D.E., Ogawa A., O'Neill K.A.,
 RA Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,
 RA Ujjainwalla F., Altman S.W., Chapman K.T., Thornberry N.A.;
 RT "The target of ezetimibe is Niemann-Pick C1-Like 1 (NPC1L1).";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137(2005).
 CC -!- FUNCTION: play a major role in cholesterol homeostasis. Is
 CC critical for the uptake of cholesterol across the plasma membrane
 CC of the intestinal enterocyte. Is the direct molecular target of
 CC ezetimibe, a drug that inhibits cholesterol absorption (by
 CC similarity). Lack of activity leads to multiple lipid transport
 CC defects. The protein may have a function in the transport of
 CC multiple lipids and their homeostasis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation
 CC of brush border membranes from proximal enterocytes suggests
 CC considerable association with the apical membrane fraction. Exists
 CC as a predominantly cell surface membrane expressed protein.
 CC -!- TISSUE SPECIFICITY: Small intestine showed the highest level of
 CC expression. Expression in other tissue including gall bladder,
 CC liver, testis, and stomach is also observed. Along the duodenum-
 CC ileum axis, the levels vary in different segments of intestine,
 CC with peak expression in the proximal jejunum. Protein expression
 CC is confined to the enterocyte. Discrete localization to the
 CC epithelial layer bordering the luminal space along the crypt-
 CC villus axis. Protein expression in the enterocyte is observed
 CC closest to the luminal space. Expression in enterocytes from the
 CC proximal (jejunum) but not in the distal (ileum) region.
 CC -!- PTM: Highly glycosylated.
 CC -!- SIMILARITY: Belongs to the patched family.
 CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AY437867; AAR97888.1; -; mRNA.
 DR RGD; 1303135; Npc1l1.
 DR InterPro; IPR004765; NP_C type.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD 5TM.
 DR Pfam; PF02460; Patched; 1.
 DR TIGRFAMs; TIGR00917; 2A060601; 1.
 DR PROSITE; PS50156; SSD; 1.
 KW Cholesterol metabolism; Glycoprotein; Lipid metabolism; Signal;
 KW Steroid metabolism; Transmembrane.
 FT SIGNAL 1 20 Potential.
 FT CHAIN 21 1331
 FT TOPO_DOM 21 282 Extracellular (Potential).
 FT TRANSMEM 283 303 1 (Potential).
 FT TOPO_DOM 304 352 Cytoplasmic (Potential).
 FT TRANSMEM 353 373 2 (Potential).
 FT TOPO_DOM 374 632 Extracellular (Potential).
 FT TRANSMEM 633 653 3 (Potential).

FT TOPO_DOM 654 665
 FT TRANSMEM 666 686
 FT TOPO_DOM 687 696
 FT TRANSMEM 697 717
 FT TOPO_DOM 718 742
 FT TRANSMEM 743 763
 FT TOPO_DOM 764 776
 FT TRANSMEM 777 797
 FT TOPO_DOM 798 846
 FT TRANSMEM 847 867
 FT TOPO_DOM 868 1113
 FT TRANSMEM 1114 1134
 FT TOPO_DOM 1135 1142
 FT TRANSMEM 1143 1163
 FT TOPO_DOM 1164 1165
 FT TRANSMEM 1166 1186
 FT TOPO_DOM 1187 1206
 FT TRANSMEM 1207 1227
 FT TOPO_DOM 1228 1242
 FT TRANSMEM 1243 1263
 FT TOPO_DOM 1264 1331
 FT DOMAIN 632 797
 FT CARBOHYD 53 53
 FT CARBOHYD 85 85
 FT CARBOHYD 138 138
 FT CARBOHYD 244 244
 FT CARBOHYD 416 416
 FT CARBOHYD 431 431
 FT CARBOHYD 464 464
 FT CARBOHYD 479 479
 FT CARBOHYD 497 497
 FT CARBOHYD 506 506
 FT CARBOHYD 606 606
 FT CARBOHYD 626 626
 FT CARBOHYD 909 909
 FT CARBOHYD 917 917
 FT CARBOHYD 996 996
 FT CARBOHYD 1038 1038
 FT CARBOHYD 1076 1076
 FT CARBOHYD 1331 AA; 146415 MW; 2E10EF2E3A337F70 CRC64;
 SQ SEQUENCE

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
 Best Local Similarity 77.8%; Pred. No. 0;
 Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRCGLWALLLLRLAQSEPVYTHIQPGYCAFYDECGKNPELSGLSLTSLNSVCLSN 60
 DB 1 MAEAGLRCGLWALLLLRLAQSEPVYTHIQPGYCAFYDECGKNPELSGLSLTSLNSVCLSN 59
 QY 61 TPARKITGDHLILQKICPRLYTGPNTQ-ACCSAKQLVSLBASISITKALLTRCPACSDN 119
 DB 60 TPARKITGDHLILQKICPRLYTGPNTQ-ACCSAKQLVSLBASISITKALLTRCPACSDN 119
 QY 120 FVNLCHNTCPNQLSFLINVTVAQLGAGQIPAVVAEAFYQHSFAEQSYDSCSRVPA 179
 DB 120 FVNLCHNTCPNQLSFLINVTVAQLGAGQIPAVVAEAFYQHSFAEQSYDSCSRVPA 179
 QY 120 FVSLCHNTCPDQSLFINTVVRVVERGAGEPPAVVAEAFYQHSFAEQSYDSCSRVPA 179
 DB 120 FVSLCHNTCPDQSLFINTVVRVVERGAGEPPAVVAEAFYQHSFAEQSYDSCSRVPA 179
 QY 180 AATLAVGTMCVGYGALCNQRLNFGQDTGNGLAPLDITFHLLPFGQAVSGIQLNEG 239
 DB 180 AATLAVGTMCVGYGALCNQRLNFGQDTGNGLAPLDITFHLLPFGQAVSGIQLNEG 239
 QY 240 VARNCSQGDVATCSCDCAASCPAIARPAQALDSTFYLGQMPGSLVLIILCSFPAVT 299
 DB 240 VARNCSQGDVATCSCDCAASCPAIARPAQALDSTFYLGQMPGSLVLIILCSFPAVT 299
 QY 240 IAPCNESQGDSDAVSCQDCAASCPVIPPPEALRPSFYMGPMGWLIIITFAVFLIS 299
 DB 240 IAPCNESQGDSDAVSCQDCAASCPVIPPPEALRPSFYMGPMGWLIIITFAVFLIS 299
 QY 300 ILLVGFVPAPADKSKMDVKKGTSLSDKLSFSTHTLLGQFGQGTWVASWPLTILVLS 359
 DB 300 ILLVGFVPAPADKSKMDVKKGTSLSDKLSFSTHTLLGQFGQGTWVASWPLTILVLS 359
 QY 360 VIVPVALAAGLVFTLTTPVLMWSAPNSQARSEKAFHQHFGFPFRNTQVLTAPNRSS 419
 DB 360 VIVPVALAAGLVFTLTTPVLMWSAPNSQARSEKAFHQHFGFPFRNTQVLTAPNRSS 419

DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Niemann-Pick C1.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
 OC Felis.
 OC NCBI_TaxID=9685;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22695580; PubMed=12809639; DOI=10.1016/S1096-7192(03)00074-X;
 RA Somers K.L., Royals M.A., Carstea E.D., Rafi M.A., Wenger D.A.,
 RA Thrall M.A.;
 RT "Mutation analysis of feline Niemann-Pick C1 disease."
 RL Mol. Genet. Metab. 79:99-103 (2003).
 DR EMBL; AF503634; AAM27451.1; -; mRNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008158; F:hedhog receptor activity; IEA.
 DR GO; GO:0030301; P:cholesterol transport; IEA.
 DR InterPro; IPR004765; NP C type.
 DR InterPro; IPR003392; patched.
 DR InterPro; IPR000731; SSD 5TM.
 DR Pfam; PF02460; Patched; I.
 DR TIGRFAMs; TIGR00917; 2A060601; 1.
 DR PROSITE; PS50156; SSD; 1.
 SQ SEQUENCE 1276 AA; 141710 MW; C2DD3496472B71EF CRC64;

 Query Match 34.8%; Score 2405; DB 2; Length 1276;
 Best Local Similarity 40.7%; Pred. No. 3.8e-161; Indels 86; Gaps 29;
 Matches 535; Conservative 232; Mismatches 461;

 QY 7 RGWLLWALLRLRLAQSPPYTHIQPGYCAFYDCGKNGPELGSGLMTLSNVCSLNTPKARKI 66
 DB 4 RGPALGLLLLLLC---PAQVLAQS---CIWYEGC---IASGD---KRYNCKYSGPKPL 51

 QY 67 TGDHLILLOKICPLRYTGNTQAQCSAKOLVSLASLSITKALLTRCPACSDNFNLHCH 126
 DB 52 PKDGYDLVQELCPGFPF-DNVSLLCCDVQLOTLKONLQLFLQSLRCPSCFYNLNLFCE 110

 QY 127 NTCSPNQSLFINVTR---VAOLGAGQLPAVVAYEAFYQHSFAEQSDYSCSRVRPAAT 182
 DB 111 LTCSPQSGFLNVTATEDVDPTNQTKNVXELQYIGESFANAMYNACRDVEAPSSND 170

 QY 183 LAVGTGCVGYGSAALCNAQRLNFGSDTGNGLAPLDIT--PHLLEPGQAVGSGIQLNEGV 240
 DB 171 KALGLLCGKDAEA-CNATNWIETMFSKNGQAPFTITPIFSDLPT-----HGMEPNNAT 224

 QY 241 ARCNSQGDVATCSQDCQDCAASC-----PAIARQALDSTFVLGMPGSLVLIIL 291
 DB 225 KGCDSEVDEVTGPCSCQDCSIVCGPKQPPPPPPVPRILGLDAMVIMWITMAFLVFF 284

 QY 292 CSVFA-----VVTILLGPRVA---PARDKSNVDPKKGTSLSLXSFSTHTLLGQ 339
 DB 285 GAFFALWCYKRYFVSEYPTIDNSIAFSVNANDRG-----ASCCDALGAFAEGLRR 337

 QY 340 FFGWGWTWASWPLTILVLSIPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHQ 399
 DB 338 LFSQWGSFCVRNPGPIIFSLAFIAACSSGLVFRVTTNPVDLWSAPSSQARLEKEYFT 397

 QY 400 HFGPFRTRNQVILITAPNRSSRYRDSLLGPK-NFSGILDLDLLELLEQLERLHLQWS 458
 DB 398 HFGPFRTRTEQLIIQAPHTSAHTQYPSGSDVPFGPPLDLAILHQLVDLQTAIEN--ITA 455

 QY 459 PEAQRLISLDQICVAPLNPNNTSLYCCINSLIQYFQNNRTLLLTANQTLMGQTSQV-- 516
 DB 456 SYNNEVTLLQDICIAPLSFYNK---NCTILSVLNYFQNSHSM-----DHEIGDDFFVA 507

 QY 517 DWKDHFLYCANAPLTKDGTALALSCADYGAVPFPFLAIGGYKGKDYSEAEALIMTFLS 576
 DB 508 DYHTHLILCYVRAPASLNDTSLLDHPCIGTFGGVFPFWLVGGYDQNNYNATALVITTFV 567

QY 577 NNYPAGDPRLAQAKLWEEAFLEBMRAFORMMACMFOVTFTAERSLEDEINRTAEDLPF 636
 DB 568 NNYNTDTERLQKARHWEKEFINFKYKN---PNLTISFTTERRSIEDLNRENGDIPV 624

 QY 637 ATSYVIFLYISIALGSYSWSRVMDSKATLGLGGVAVVLGAVMAAMGPFSLGTRSSL 696
 DB 625 IISVAINMFLYISIALGHIKSCSRLVDSKISLGIAGILIVLSSVACSLGIFSYVIGIPLT 684

 QY 697 VILQVPFLVLSVADNIFIFVLEYQRLPRPCEPREVHIGRALGRVAPSMILCLSEAI 756
 DB 685 IVIEIPLFLAVGVNDNIFILVQTYQDERLHGETLDQQLGRVLGEVAPSMFLSSFEAV 744

 QY 757 CFELGALTMPAVRTPALTSGLAVILDFLQMAFAVALLSDSKRQBSASRLDVCCKVPK 816
 DB 745 AFLGALSKMPAVHTSFLPAGMAVILDFLQIYCFVSLGLDICKRQENRDLVLCVGRS 804

 QY 817 ELPPPGQ-GEGLLGGFFQKAYAPPLHWTIRGVLLFLALFGVLSYMSCHISVGLDQEL 875
 DB 805 EDGTSVQASESCLFRLFKHSYSPLLKDMRPVIAIFVGLSFSVAVLNKRVEIGLQSL 864

 QY 876 ALPKDSYLLDYFLFNRYFVGAPVYFVTTLGYNFSSEAGNNAICSSAGCNGNPFSTOKIQ 935
 DB 865 SMPDSDSYMDYFKSL-KYLHAGPPVYFVLEBEGHDYTSLKQGNMVGCGMCGNNDLSVQOIF 923

 QY 936 YATFPEQSYLATIPASSWDDFDLWLTTP-SGCCRLYISGPNKOKFCPSTVNSLNLKNCM 994
 DB 924 NNAQDSYTRIGPAPSSWIDDYDFWVKPQSSCCRVNS---TDRFCNASVVDPAICR-CR 979

 QY 995 SITM-GSVRPSVEQFHKYLPWFLNDRNINIKCPKGLAAYSTSVN-LFSDGOVLASRWAY 1052
 DB 980 PLTQEGKQRPQGGDFMFLPMFLSDNPNPKCGKGHAAYSSAVNINILGNDTGVGATYFTY 1039

 QY 1053 HKPLKNSQDYTEALRAARELAANITADLRKVPGTDP-EEVFPYTTITNVFYEQLTILPE 1111
 DB 1040 HTVLQTSADFTDMARKANLITASNIT---KTMGLEGSNRYRVPFSVVFVEQYLTID 1095

 QY 1112 GLFWLSCLVPTFAVSCLLLGLDLRLSGLNLLSTVMIIVDTVGFPMALWDISYNVSLNL 1171
 DB 1096 TIFNLVSLGALFLVTVILLGCDLWSAVIMCITIAMILVNMFGVWMLWGISLNAVSLNL 1155

 QY 1172 VSAGMSEVFVSHITRFAISTKPTWLERAKEATISGSAVPAGVAMTNLPGILVLGLAK 1231
 DB 1156 VMSCGISVEFCSHITRAFTVSMKGSRAQRAEALAHMGSSVFSGITLTKFGIVVLAPAK 1215

 QY 1232 AOLIQIFPRLNLTITLGLLHGLVFLPVLISYVGGPDVNPALAL---EQKRAEE 1282
 DB 1216 SQIFQIFRMYLAWVLGLGATHGLIFLPVLISYIGPSINKAKSLATQEQYKGT 1269

 RESULT 8
 Q9GK52 CANPA
 ID Q9GK52 CANPA PRELIMINARY; PRT; 1276 AA.
 AC Q9GK52;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Niemann-Pick type C1 disease protein.
 GN Name=NFC1;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Murakawa M., Freeman M.W.;
 RA Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF315034; AAG40873.1; -; mRNA.
 DR Ensembl; ENSCAFG0000018183; Canis familiaris.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0016020; C:membrane; IEA.

DR GO: 0008158; P.hedgehog receptor activity; IEA.

DR GO: 0003001; P.cholesterol transporter; IEA.

DR InterPro: IPR004765; NP_C type.

DR InterPro: IPR003392; Patched.

DR InterPro: IPR000731; SSD 5TM.

DR Pfam: PF02460; Patched; 1.

DR TIGRFAMs: TIGR00917; 2A060601; 1.

DR PROSITE: P050156; SSD; 1.

SQ SEQUENCE 1276 AA; 141695 MW; 07D1327B51BD5681 CRC64;

Query Match 34.8%; Score 2404.5; DB 2; Length 1276;

Best Local Similarity 40.4%; Pred. No. 4.2e-161;

Matches 524; Conservative 233; Mismatches 434; Indels 105; Gaps 28;

QY 33 CAFYDECGRNPELSGLMTLSNVSCLSNTPARKITGDHLLILQKICPRLYTGENTQACCS 92

DB 25 CVWYGECC--IASGD---KRYNCQYSGPPKPLPKQGYDLAQELCPGLFF-DNVSVCDD 76

QY 93 AKQLVSLASLITKALLTRCPACSDNFVNLHCHTCSPNQSPLFNTR-----VAQLGAG 148

DB 77 VQQLRLKDSQLPLQLFLSRCPSCFYNLMNLFCELTCSPPSQQLFNVTTEDYDVPVTNQ 136

QY 149 QLPVAVYAEAFQHSFAEQSDYSCSRVRVPAATLAVGTMCVGYSGALCNAQRLNFGQD 208

DB 137 TKTNVKELQYVYGESFANAMYNACRDVEAPSSNDKALGLLCKGAEAE-CNATNWLEYMFN 195

QY 209 TNGGLAPL-----DITPHLLEPGQAVSGIGIQLNEGVARCNESQGDVATCSQDCCAA 261

DB 196 KNGQAPFTIIFSDLPALH-----GKPMNATKGCDEPDEVETAPCSQDCSV 245

QY 262 SC-----PAIRAPQALDSTFYLGQMPGSLVLIILCSVPVAVTILLVGRVAPARD 312

DB 246 VCGPKPQPPAPAPWRILGLDAMVIMWTYMAFLMPPGAFPAVWCY----- 293

QY 313 KSKWVDPKGTSLDKLSSTHT-----LLGQPFQ-----WGTVASWPLT 354

DB 294 -RKRYFVSEYTPIDNSIAFVNAGDTGEASCCDALGAEPGCLRLFTQWGSCFIRNPGC 352

QY 355 ILVLSVIPVALAAGLVFELTTDPVELWSAPNSQARSEKAFDQHPGFPFRTNVLTA 414

DB 353 IIFSLAPTAACSGGLVFSRVTPNVDLWSARGSQARLEKYPDAHFGFPFTEQLIQA 412

QY 415 PNRSRYDSSLGPK-NPFGILDLDLLELELEQLERHLQWSPQAQRNLSQDICVA 473

DB 413 PHTSVHTQYPYSGSDVPPGPELDIGILHVLQDLQTAIEN--ITATYNNETVTLQDICVA 470

QY 474 PLNPNTSLYDCINSLLQYFQNNRTLLLTANQTLMGTSQV--DWKDHFLYCANAPLT 531

DB 471 PLSPYNNK---NCTIMSVLNYFQNSHML-----DHKIGDDFYVYADYTHLLYCVRAPAS 522

QY 532 PKDGTALALSCMADYGAPVFPFLAIGYKGDYSBAELIMTFSLNNYPAGDPRLAQAKL 591

DB 523 LNDTSLHDPCLGTGGVFPVPLVGGYDDQNNATALVITFPVNNYNTDEKLRQAQ 582

QY 592 WEEAFLEENRAFQRRMAGMFQVFTAERSLEDEINRTAEDLPIFATSYVIFLYLSIAL 651

DB 583 WEKEFINFKVYEN---PNLTISFTTERSIEBELNRESNGDVFTVLISAVNMFYLSIAL 639

QY 652 GYSWSRVMVDSKATLGLGGVAVVLGAVMAAGFPYSLGIRSSILVLOVPLVLSVGA 711

DB 640 GHKCSRLPVDKSLKISLAGILVLSVCMCSLIGIFSYFGIPLTIVIVIPFLVAVGV 699

QY 712 DNIETFLVLEYQRLPRPGEPRVHLGRALGRVAPSMLLCSLSEACFCFLGALTPPAVRT 771

DB 700 DNIETFLVLEYQRLPRPGEPRVHLGRALGRVAPSMFLSSFSSEAVFAFFGALSQMPAHT 759

QY 772 FALTGLAVILDQLQMSAFVALLSDSKRQBSRLDVCCKVKPQELPPGQG----EGL 827

DB 760 FSLPAGMAVLIDFLQITCFVSLGLDLKQEKRLDVLCLLTGSE---GGTGIQASESC 816

QY 828 LIGFTQKAYAPFLHWTIGVVLVLLFLALFGVSLYSMSCHISVGLQDELAKPDSYLLDYF 887

DB 817 LPRFFKNSYSPFLKDWMRPIVIAVFVGLSFLSIAVLNKNVETGLQSLSPDSDSYMDF 876

QY 888 LFLNRYFEVGAPVYFVTTLTGYNFSEAGMNAICSSAGCNFNFTQKIQIYATEPPEQSYLA 947

DB 877 KSL-KYLHAGPPVYFVLEBHDYTSLEGQNMVCGGCMNDLSLVQIPSAQLDNYTHIG 935

QY 948 IPASSWDDFDWLTP-SSCCRLYISGPNKDFCFSTVNSLKNCKMCSITM-GSVRPSV 1005

DB 936 PAPSSWIDDDYDFWVKPQSSCCRVYNS---TDQFCNASVVDPAACVR-CRPLTQEGKRRPQ 991

QY 1006 EOPHXYLPWFNDRPNIKCPKGLAAYSTSVNLT-SDGOVLASRFMAYHKPKNSODYTE 1064

DB 992 EDFMFLPWFNLSNPNPKCKGCHAYGSAVNLVGNDSVSGATYFWYTYTTLQTSADFTD 1051

QY 1065 ALRAARELAANITADLRKVPGTDPAPFVPPYITITVTFYEQYLTIPEGLFMLSCLVPTTF 1124

DB 1052 AMRKAILIASNTKTM-SLKGSN--YRVFPYFVYFVYEQYLTIIDDTFNLVSVSLGAIF 1108

QY 1125 AVSCILLGLDLBSGLNLLSIYVNLVDVTFGMALWDISVNLVSLNLSVAVGMSVBFVSH 1184

DB 1109 LVTVLGCELMSAVIMCVTIAMILVNMFMGLAGISLNVSLVNLVMSGSGISVBFCSH 1168

QY 1185 ITRSPAISTKPTMLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFPRLNL 1244

DB 1169 ITRATVSAKGRVERAREALSHMGSSVFSGITLTKFGIVVLAFAKSIQFVFTITGMYL 1228

QY 1245 LITLGLHLGLVFLPVLISYVGPVNPALAL-EQKR 1279

DB 1229 AMVLGATHGLIFLPLVLLSYIGPSINKAKSLASQER 1264

RESULT 9

NPCL HUMAN

ID NPCL HUMAN STANDARD; PRT; 1278 AA.

AC O15118; Q9P130;

DT 30-MAY-2000 (Rel. 39, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Niemann-Pick C1 protein precursor.

GN Name=NPCL;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE, VARIANT ILE-642, AND VARIANTS NPCL

RX MEDLINE=97362323; PubMed=9211849; DOI=10.1126/science.277.5323.228;

RA Carstee B.D., Morris J.A., Coleman K.G., Loftus S.K., Zhang D.,

RA Cummings C., Gu J., Rosenfeld M.A., Pavan W.J., Krizman D.B.,

RA Nagle J., Polymeropoulos M.H., Sturley S.L., Ioannou Y.A.,

RA Higgins M.B., Conly M., Cooney A., Brown A., Kaneski C.R.,

RA Blanchette-Mackie E.J., Dwyer N.K., Neufeld E.B., Chang T.-Y.,

RA Lisicum L., Strauss J.F. III, Ohno K., Zeigler M., Carmi R., Sokol J.,

RA Markie D., O'Neill R.R., van Diggelen O.P., Ellender M.,

RA Patterson M.C., Brady R.O., Vanier M.T., Pentchev P.G., Tagle D.A.;

RT "Niemann-Pick C1 disease gene: homology to mediators of cholesterol

RT homeostasis."

RL Science 277:228-231 (1997).

RN [2]

RP NUCLEOTIDE SEQUENCE, AND VARIANTS.

RX MEDLINE=99355599; PubMed=10425213; DOI=10.1006/bbrc.1999.1070;

RA Morris J.A., Zhang D., Coleman K.G., Nagle J., Pentchev P.G.,

RA Carstee B.D.;

RT "The genomic organization and polymorphism analysis of the human

RT Niemann-Pick C1 gene."

RL Biochem. Biophys. Res. Commun. 261:493-498 (1999).

RN [3]

RP NUCLEOTIDE SEQUENCE, AND VARIANT ILE-642.

RX MEDLINE=21623216; PubMed=11754101; DOI=10.1002/humu.10016;

RA Bauer P., Kuestermann R., Bauer C., Finckh U., Hufen A., Kropp J.,

RA Braun S., Kuestermann-Kuhn B., Schmidt D., Harzer K., Rolfs J.,

RT "NPCL: complete genomic sequence, mutation analysis, and

RT characterization of haplotypes."

RL	Hum. Mutat. 19:30-38 (2002).
RP	[4]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT IL6-642.
RC	TISSUE=Placenta;
RA	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altshuler S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley C.K., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodríguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA	Schuerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RL	[5]
RP	CHARACTERIZATION.
RP	MEDLINE=99128318; PubMed=9927649; DOI=10.1073/pnas.96.3.805;
RA	Watarai H., Blanchette-Mackie E.J., Dwyer N.K., Glick J.M., Patel S.,
RA	Neufeld E.B., Brady R.O., Pentchev P.G., Straus J.F. III;
RT	"Niemann-Pick C1 protein: obligatory roles for N-terminal domains and
RT	lysosomal targeting in cholesterol mobilization.";
RL	Proc. Natl. Acad. Sci. U.S.A. 96:805-810 (1999).
RL	[6]
RP	VARIANT NPD TRP-992.
RP	MEDLINE=98299797; PubMed=9634529;
RA	Greer W.L., Riddell D.C., Gillan T.L., Girouard G.S., Sparrow S.M.,
RA	Byers D.M., Dobson M.J., Neumann P.E.;
RT	"The Nova Scotia (type D) form of Niemann-Pick disease is caused by a
RT	G309T-->T transversion in NPC1.";
RL	Am. J. Hum. Genet. 63:52-54 (1998).
RL	[7]
RP	VARIANTS NPC1 GLN-934; LEU-940; ASN-948; LEU-954; TRP-992; ALA-1007;
RP	THR-1061 AND VAL-1213
RA	MEDLINE=99452586; PubMed=10521290;
RA	Greer W.L., Dobson M.J., Girouard G.S., Byers D.M., Riddell D.C.,
RA	Neumann P.E.;
RT	"Mutations in NPC1 highlight a conserved NPC1-specific cysteine-rich
RT	domain.";
RL	Am. J. Hum. Genet. 65:1252-1260 (1999).
RL	[8]
RP	VARIANT NPC1 THR-1061.
RP	MEDLINE=99452593; PubMed=10521297;
RA	Millat G., Marçais C., Rafi M.A., Yamamoto T., Morris J.A.,
RA	Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;
RT	"Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant
RT	allele in patients of Western European descent and correlates with a
RT	classic juvenile phenotype.";
RL	Am. J. Hum. Genet. 65:1321-1329 (1999).
RL	[9]
RP	VARIANTS NPC1, AND VARIANTS ARG-215; VAL-858 AND GLN-1266.
RP	MEDLINE=99408226; PubMed=10480349; DOI=10.1007/s004300051057;
RA	Yamamoto T., Namba E., Ninomiya H., Higaki K., Taniguchi M., Zhang H.,
RA	Akashi S., Watanabe Y., Takehima T., Inui K., Okada S., Tanaka A.,
RA	Sakuragawa N., Millat G., Vanier M.T., Morris J.A., Pentchev P.G.,
RA	Ohno K.;
RT	"NPC1 gene mutations in Japanese patients with Niemann-Pick disease
RT	type C.";
RL	Hum. Genet. 105:10-16 (1999).
RL	[10]
RP	VARIANTS NPC1 GLN-958 AND ALA-1007.
RP	MEDLINE=2131311; PubMed=11349231;
RA	Millat G., Marçais C., Rafi M.A., Yamamoto T., Morris J.A.,
RA	Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;
RT	"Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant
RT	allele in patients of Western European descent and correlates with a
RT	classic juvenile phenotype.";
RL	Am. J. Hum. Genet. 65:1321-1329 (1999).
RL	[9]
RP	VARIANTS NPC1, AND VARIANTS ARG-215; VAL-858 AND GLN-1266.
RP	MEDLINE=99408226; PubMed=10480349; DOI=10.1007/s004300051057;
RA	Yamamoto T., Namba E., Ninomiya H., Higaki K., Taniguchi M., Zhang H.,
RA	Akashi S., Watanabe Y., Takehima T., Inui K., Okada S., Tanaka A.,
RA	Sakuragawa N., Millat G., Vanier M.T., Morris J.A., Pentchev P.G.,
RA	Ohno K.;
RT	"NPC1 gene mutations in Japanese patients with Niemann-Pick disease
RT	type C.";
RL	Hum. Genet. 105:10-16 (1999).
RL	[10]
RP	VARIANTS NPC1 GLN-958 AND ALA-1007.
RP	MEDLINE=2131311; PubMed=11349231;
RA	Millat G., Marçais C., Rafi M.A., Yamamoto T., Morris J.A.,
RA	Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;
RT	"Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant
RT	allele in patients of Western European descent and correlates with a
RT	classic juvenile phenotype.";
RL	Am. J. Hum. Genet. 65:1321-1329 (1999).
RL	[9]
RP	VARIANTS NPC1, AND VARIANTS ARG-215; VAL-858 AND GLN-1266.
RP	MEDLINE=99408226; PubMed=10480349; DOI=10.1007/s004300051057;
RA	Yamamoto T., Namba E., Ninomiya H., Higaki K., Taniguchi M., Zhang H.,
RA	Akashi S., Watanabe Y., Takehima T., Inui K., Okada S., Tanaka A.,
RA	Sakuragawa N., Millat G., Vanier M.T., Morris J.A., Pentchev P.G.,
RA	Ohno K.;
RT	"NPC1 gene mutations in Japanese patients with Niemann-Pick disease
RT	type C.";
RL	Hum. Genet. 105:10-16 (1999).
RL	[10]
RP	VARIANTS NPC1 GLN-958 AND ALA-1007.
RP	MEDLINE=2131311; PubMed=11349231;
RA	Millat G., Marçais C., Rafi M.A., Yamamoto T., Morris J.A.,
RA	Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;
RT	"Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant
RT	allele in patients of Western European descent and correlates with a
RT	classic juvenile phenotype.";
RL	Am. J. Hum. Genet. 65:1321-1329 (1999).

Qy	164	PAE0SYDSCSRVRPAAATLAVGTWCGVYGSALCNAQWLNFGQDGTGGLAPLDIT	219
Db	163	FANAMYNAACRDVEAPSSDKALGLCGKQADA-CNATNMIBYMFNKONGQAPFTTIPVFS	221
Qy	220	-PHLLEPQAVGSGIOPLENEGVARNESQGDVATCSQDCAASCAPAIARPQ-----	270
Db	222	DPFV-----RGMEPNNAATKGDSESDVETAPCSQDCSIVCGPKPQPPPPAPWTI	273
Qy	271	-ALDSTFYLGMPGSLVLIILCSVFAVVITLL-----VGFRVAPARDKSKM	316
Db	274	LGLDAMYIMITYMAELVFFGAPFAVVCYRKRVFVSEYTPIDSNIAFSV-NASDKGE-	331
Qy	317	VDPKGTSLSKLSFSTHTLQGFQFGWGTWASWPLTILVSVIPVVALAAGLVFTLT	376
Db	332	-----ASCDDPVSAAPEGLRLRTRWGSFCVRNPGCVIFPSLVFITACSGLVFVRVT	385
Qy	377	TDPELWAPNSQARSEKAFDQHFQPFRTNQVILTAPNRSSVRYDSLLGPK-NFSGI	435
Db	386	TNPVDLWAPSQAARLEKEYDQHFQPFRTNQVILTAPNRSSVRYDSLLGPK-NFSGI	445
Qy	436	LDLDLLLELLEQLERHLQWSPBAQRNLSLQDICYAPLNDPNTSLYDCINSLLQVFO	495
Db	446	LDIQLHQLVLDLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVLYNFO	500
Qy	496	NRTTLLLTANQTLMGQTSQDWDKDHLYCANAPLTPKDGTAALSQWADYGAPVPEPLA	555
Db	501	NSHSLVDHKKGDFFF---VYADYTHFLYCVNRPASLNDTSLHDPCLGTGGPVFPLV	557
Qy	556	IGCYKGYSEALIMTFSNYPAGOPRLAQAALWEEAFLEENRAPQRMRMAGMFOVTF	615
Db	558	LGYYDDQYNNATLVIPVNNYNDTEKQBAWEKEFINFKYKN---PNLTISF	614
Qy	616	TAERLEDEBINTTADLPFATSVIPLFYLSLALGYSVSRVWVDSKATLGLGVAV	675
Db	615	TAERSIEDELNESDSDVFTVSVIAPLFIYLSLALGYSVSRVWVDSKATLGLGVAV	674
Qy	676	VLGAVWAMGFSYLGIRSLVILQVPELVLSVGDADNIFIVLSVQRLPRPGRPREVH	735
Db	675	VLSVACSGVFSYGLPLTLVIEVPLVLAVGVDNIFILVQAYQDERLQGTTLQOQ	734
Qy	736	IGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALL	795
Db	735	LGRLVGEVAPSMFLSFSFETVAFGLSVMVAHVTFSLFAGLAVFIDFLQITCFVSL	794
Qy	796	SLDSKQKQASRLDVCVCPQELPPRQ--GEGLLGFFQKAYAPFLHWTGCVLLFL	854
Db	795	GLDIKQKRNLDIFCCVRGAEVGTSVQASESCLFRFFKNSYSLLLKDMRPIVIAIPV	854
Qy	855	ALFPGVLSYNSCHTSVGLDQELALPKDSYLLDYFLNRYFVGVAPVYFVTTTLGYNFSSEA	914
Db	855	GVLSFSAVLNKLVDIGLQSLGMPDDSDYVDFVFKSISQYLHAGPPYFVLEGHDTSK	914
Qy	915	GMNAICSSAGCNFSTQKIQVATEPFEQSILAIPASSWVDVDFDMLTP-SSCCRYISG	973
Db	915	QGNWVCGGNCNDSLVIQIFNAALQDNTVTRIGFAPSSMIDDYFDWVKPQSSCCRV--D	971
Qy	974	PNKDFCPSTVNSLNCNKMSIT-NGSVRPSVEQHKYLPWFLNDRPNIKCPKGLAAY	1032
Db	972	NITDQFCNASVDPACVR-CRPLTPGKQRPQGGDFMRFLPMLSDNPNPKCGKGHAAY	1030
Qy	1033	STSVN--LTSQGVLASRWAYHKPLKNSQDYTEALRAARELANATADLRKVPVGDPAF	1090
Db	1031	SSAVNLLHGTRGATYFTVTVVQTSADFDALKAELIASNTV-ETWINGS--AY	1087
Qy	1091	EVFPYITNPFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDRLSGLLNLISVIMLV	1150
Db	1088	RVPFYSVFYFYEQYLTIIIDTIFNLGVSLSGAFVLTWLLGCELASVIMCATIAMLV	1147
Qy	1151	DTVGFMALNDISYNAVSLINLVAGVNSVFVSHITFSFAISKPTWLERAKEATISMG	1210
Db	1148	NMFQVWMLWGISLNAVSLNVLVMSGIGSVBFCSHITRAFTVSMKGRVREAREALAHMG	1207

Qy	1211	AVFAGVAMTNPLGILVLGLAKAQLIQIPFRFLNLLITLLGLLHGLVPLPVLVSYGPDVN	1270
Db	1208	SVFSGITLTKFGGIVVLAFKASQIFQIFRMYLAWVLGATHGLIFLPLVLLSYIGPSVN	1267
Qy	1271	PA 1272	
Db	1268	KA 1269	
RESULT 11			
Q7TMD4 MOUSE			
ID	Q7TMD4	MOUSE PRELIMINARY;	PRT; 1277 AA.
AC	Q7TMD4		
DT	01-OCT-2003	(T-EMBLrel. 25, Created)	
DT	01-OCT-2003	(T-EMBLrel. 25, Last sequence update)	
DT	10-MAY-2005	(T-EMBLrel. 30, Last annotation update)	
DE	Niemann Pick type Cl.		
GN	Name=Npci;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C57BL/6; TISSUE=Brain;		
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Uadin T.B., Tohiyuki S., Carninci P., Mullahy S.J.,		
RA	Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smal M.A.,		
RA	Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C57BL/6; TISSUE=Brain;		
RC	Strausberg R.;		
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.		
[3]			
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C57BL/6; TISSUE=Brain;		
RC	Strausberg R.;		
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC052437; AAH52437.1; -; mRNA.		
DR	EMBL; BC054539; AAH54539.1; -; mRNA.		
DR	GO; GO:0006897; P.endocytosis; IMP.		
DR	InterPro; IPR004765; NP_C-type.		
DR	InterPro; IPR003392; Patched.		
DR	InterPro; IPR000731; SSD 5TM.		
DR	Pfam; PF02460; Patched; 1.		
DR	TIGRFAMs; TIGR00917; 2A060601; 1.		
DR	PROSITE; PS50156; SSD; 1.		
SQ	SEQUENCE 1277 AA; 142882 MW; 3B42230AAC8E564E CRC64;		
Query Match 34.7%; Score 2400; DB 2; Length 1277;			
Best Local Similarity 40.4%; Pred. No. 8.7e-161;			
Matches 528; Conservative 233; Mismatches 443; Indels 104; Gaps 27;			
Qy	11 LWALLRLAQSPFYTIHQPGYCAFYDECKGNPELSSGLMTLSNVSCLSNTPARKITGDH		70

QY 577 NNYPADGPELAQKWEAFLEBEMRAFORRMAGMFOVTTAERSLEDEINRTAEDLPF 636
Db 568 NNYNDTERLQKAHWEKEFINFVKNYK---PNLTISTPSTERSDELRNSGDIPTV 624
QY 637 ATSYIVIFYISIALGYSYSWRSVMVDKATLGLGVAVVILGAVMAAMGFYSVLGRSSL 696
Db 625 IISYAIMFYISIALGHIKSRLLVDKISIGIAGILIVLSVACSLGIFSVGIPTL 684
QY 697 VILQVFPFVLSVGADNIPIFVLEYQRLPRRPEPREVHIGRALGRVAPSMLLCSLSEAI 756
Db 685 IIVIEVPFLVAVGVNDIFILVQYQDERLHGETLDQQLGRVGEVAPSMELSPSEAV 744
QY 757 CPELGTMPAVRFAITSLGAVILDLQMSAFVALLSLDKQESRLDVCCCKVQ 816
Db 745 APLGALSOMPAVHTFSLFAGMAVLIDFLQTCFVSLGLDKRQKRLNDVLCVRS 804
QY 817 ELPPPCQ-CEGLLLGPFQKAYAPFLHMTTRGVLLFLALFGLVLSYMSCHLSVGLDQSL 875
Db 805 EDGTSVQASESCLFLPKHSYSPLLLKDMWRIVTAIFGVLSFVAVLNKVEIGLDQSL 864
QY 876 ALPKDSYLDYFLFLNRYFEVGAPEVFTVTLGYNFSSEAGMNAICSSAGCNNSFTQKITQ 935
Db 865 SMPDSDSYNDYFKSL-KYLHAGFPVVFVLEEGHDYTLKGQNMVCGMCCNDSLVQQLF 923
QY 936 YATEPEQSYLAIPASSWDDFDMLTP--SSCRLYISGNPKDKPCPTVNSLNLKNCM 994
Db 924 NAAQLDSYTRIGAFSSWDDYDFMWKQSSSCRVNS---TDRFCNASVDPACIR-CR 979
QY 995 SITM-GSVRPSVEQHKYLPWFLNRPNIKCPKGLAAVSTSVN-LTSPGQVLASREWAY 1052
Db 980 PUTQEGKBPQGGDFRFRFPMFLSNPNPKCGKGGHAAVSSAVNIIIGNDTGVGATYFMTY 1039
QY 1053 HKPLKNSQDYTEALRAARELAANITADLRKVPCTDPA-PEVFPYTTITNVFYQYLTIPL 1111
Db 1040 HTVLQTSADFTAMRKANLIASNT----KTMGLEGSNRYVPFYSVFFYQYLTIIDD 1095
QY 1112 GLFMLSCLVPTFVSCLLGLDLRSGLNLNLSIVMLIVDTVGFMAWDISYNAVSLNML 1171
Db 1096 TIFNLSVSLGAIFLVTIILLGCDLMSAVIMCITIAMILVNMFGVWMLGSLNAVSLNVL 1155
QY 1172 VSAVGMSVEFVSHITRSPATSKPTWLEBAKATISMGSAVAGVAMTNLPGILVLGLAK 1231
Db 1156 VNSCGISVEFCSHITRAFTVSMKGRQAQAEALAHMGSVFSGITLTKFGIVVLAPAK 1215
QY 1232 AQLIQIFFRLNLLITLLGLHGLVPLPVILSVGPDVNPALAL---EOKRAEE 1282
Db 1216 SQIFQIFRMYLAWVLLGATHGLFLPLVLSYIGPSINKAKSLATQEQYKGT 1269
RESULT 14
NPCL MOUSE
AC O35604; O35605; STANDARD; PRT; 1278 AA.
DT 30-MAY-2000. (Rel. 39, Created)
DT 10-MAY-2000. (Rel. 39, Last sequence update)
DT 10-MAY-2005. (Rel. 47, Last annotation update)
DE Niemann-Pick C1 protein precursor.
GN Name=Npcl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97362324; PubMed=9211850; DOI=10.1126/science.277.5323.232;
RA Loftus S.K., Morris J.A., Carstee E.D., Gu J.Z., Cummings C.,
RA Brown A., Ellison J., Ohno K., Rosenfeld M.A., Tagle D.A.,
RA Pentchev P.G., Pavan W.J.;
ET "Murine model of Niemann-Pick C disease: mutation in a cholesterol
homeostasis gene.";
RL Science 277:232-235 (1997).
RN [2]

RP SUBCELLULAR LOCATION.
RX MEDLINE=99145606; PubMed=9990080; DOI=10.1073/pnas.96.4.1657;
RA Patel S.C., Suresh S., Kumar U., Hu C.Y., Cooney A.,
RA Blanchette-Mackie E.J., Neufeld E.B., Patel R.C., Brady R.O.,
RA Patel Y.C., Pentchev P.G., Ong W.-Y.;
RT "Localization of Niemann-Pick C1 protein in astrocytes: implications
for neuronal degeneration in Niemann-Pick type C disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1657-1662 (1999).
CC -!- FUNCTION: Involved in the intracellular trafficking of
cholesterol. May play a role in vesicular trafficking in glia, a
process that may be crucial for maintaining the structural and
functional integrity of nerve terminals.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Found in late
endosomes and lysosomes.
CC -!- TISSUE SPECIFICITY: Expressed predominantly in perisynaptic
astrocytic glial processes. Also expressed in heart, spleen, lung,
liver, skeletal muscle, kidney, testis.
CC -!- INDUCTION: Activated by the drugs progesterone and U-18666A which
block cholesterol transport out of lysosomes and by the
lysosomotropic agent NH4Cl.
CC -!- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain
containing a di-leucine motif necessary for lysosomal targeting
are critical for mobilization of cholesterol from lysosomes.
CC -!- SIMILARITY: Belongs to the patched family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

CC EMBL; AF003349; AAB63372.1; -; mRNA.
CC EMBL; AF003349; AAB63373.1; -; Genomic_DNA.
CC PIR; T30188; T30188.
CC Ensembl; ENSMUSG00000024413; Mus musculus.
CC MGI; MGI:1097712; Npcl.
CC GO; GO:0006897; P: endocytosis; IMP.
CC InterPro; IPR004765; NP_C type.
CC InterPro; IPR003392; Patched.
CC InterPro; IPR000731; SSD 5TM.
CC Pfam; PF02460; Patched; 1.
CC PROSITE; PS00917; 2A060601; 1.
CC PROSITE; PS0156; SSD; 1.
KW Glycoprotein; Lysosome; Signal; Transmembrane.
FT SIGNAL 1 23 Potential.
FT CHAIN 24 1278 Niemann-Pick C1 protein.
FT TRANSMEM 271 291 Potential.
FT TRANSMEM 352 372 Potential.
FT TRANSMEM 623 643 Potential.
FT TRANSMEM 655 675 Potential.
FT TRANSMEM 685 705 Potential.
FT TRANSMEM 761 781 Potential.
FT TRANSMEM 834 854 Potential.
FT TRANSMEM 1099 1119 Potential.
FT TRANSMEM 1125 1145 Potential.
FT TRANSMEM 1196 1216 Potential.
FT TRANSMEM 1228 1248 Potential.
FT DOMAIN 621 786 SSD.
FT MOTIF 1275 1278 Di-leucine motif.
FT COMEBIAS 250 260 Poly-Pro.
FT CARBOHYD 71 71 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 123 123 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 138 138 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 186 186 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 223 223 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 229 229 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 415 415 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 460 460 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 479 479 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 525 525 N-linked (GlcNAc...) (Potential).
FT CONFLICT 474 474 S -> F (in Ref. 1; AAB63373).
FT CONFLICT 479 479 N -> D (in Ref. 1; AAB63373).

Db 196 KONGQAPFTTTPVFSDLPT-----HGMEMPNNATKCCDESDVDEVTGCPSCQDCSACVGPCK 250
QY 267 ARPQ-----ALDSTFYLQMPGSLVLIILICSVFAVVTILLVG--FRVAPARDK-- 313
Db 251 PQPPPPPPVWKILGLDAMY-----VIMNSTYMAFLLVPPGAFFAVCMVKRKYF 298
QY 314 -----SKMVDPKKTSLSKLSFSHTILLGQFFQCGWGTWVASWPLTILVLVS 360
Db 299 VSEFTPIDGNI PFSINADKGGTCCDPLGAFAEHLRLFLFWGSCFVRHFGCVVFFSV 358
QY 361 IPVVALAAGLVFTELTTPVLSWAPNSQARSEKAFHQHFGFPFRFTNOVILTPNRSY 420
Db 359 AFIAACSSGLVFIQVTTDPVLMSAPGSAQLEKEYDFTHFGFPFRTEQLIIRAPHTPH 418
QY 421 RYDLSLLGPK-NFSGILDLDLLELELERLHLQVMSPEAQRNLSQDICYAPLNPON 479
Db 419 IYEPYPSGADVPPGLPVLAVNHVLQDTAIE--SITASYNNTVTLRDICYAPLSYN 476
QY 480 TSLYDCCINSILQYFONNRTLLLTANQTLMGOTSOV--DWKDHFLYCANAPLTFKDGTA 537
Db 477 ---QNCITLSVLNYFQNSHVL-----DHQVGDDEFFYADYTHFLYVWRAPASLMDTSL 528
QY 538 LALSCHADYGAPEPPLAIGGVKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEPL 597
Db 529 LHPCLGTEGGVPPVPLVGGDDQNNATALLVITFPVNNYNTDEKLRQAQWKEKFI 588
QY 598 EEMRAFQRMAGMFOVTTAERSLEDEINRTAEDLPIFATSVIVIFYLSIALAGSYSSW 657
Db 589 NFQVQYEN---PNLITSFKAERSIEDELNRESNDVFTVLISYGVWFLYISIALGHKSC 645
QY 658 SRMVDSKATLGGVAVVLGAVMAAGPFSYLGRSSILVILQVFPFLVLSVGAENIFIF 717
Db 646 RELLVDSKILLGAGVLI VLSPVACSLGPFYSWGPPLTILVIEVIFPLVAVGVNDIFIL 705
QY 718 VLEYQRLPRPGRPREVHIGRALGVAPSMILCSLSEALCFELGALTMPAYRTALTSG 777
Db 706 VQYQIDRLQOETLQQVGRVLGEVAPSMFJSSFAETVAFVLGGLSVMPAVHTLSLFA 765
QY 778 LAVILDFLLQMSAFVALSLDSKROEASRLDYCCCKVQBELPPP-----QOEGILL 828
Db 766 MAVLIDFLQLITCFVSLGLDIIKQEKQNDLVLCV-----GGAADDAGIQASESCL 817
QY 829 LGFFOKAYAPFLHMTIRGVILLFLALFGVSLYSNMCHISVGLDQBELALPKDSYLLDYP 888
Db 818 PRFRNSYAPLLIKDMWRPLVAVFVGLVSFSLVNLKVEIGLDQSLSPDSDSYTDYFQ 877
QY 889 FLNRYFEVCAPIYVFTLGYNFSSEAGMNAICSSAGCANNFSTQKIQYATEPBAQSYLAI 948
Db 878 SLNQYLHAGPPYFVLEBEGHDTSTKGQNMVCGGLGCNNDSLQVQVFTAAQLDSYTRIGF 937
QY 949 PASSWVDDFDLWTP--SSCCRLYISGPNKDKFCPTSVNSLNLCKNCMSIT-MGSVRPSVE 1006
Db 938 APSSWIDDYFDWVKQSSCCRIYNS---TEQCNASVNVPTCV-CRPTTPGKQRPQGA 993
QY 1007 QPHKYLPLWFLNDRPNKCPKGLAAYSTSVNLTSDG-QVLASRFMAVHKPLKNSODYTRA 1065
Db 994 DPMRFLPMLSDPNPKCPKCGKGHAAVSAAVNILDNGTRVGATYFWTYHTVLQTSADFIDA 1053
QY 1066 LRAARELANIITADLRKVBCTDPAFEVFPYTTTNVFEYQYLILPEGLPMLSLCLVPTFA 1125
Db 1054 MEKARLIASINTRTNMQQGD---HRVPFYSVFPYFEQYLTWIDDTIFNLVSLSGAIFL 1110
QY 1126 VSLCLGLDLRLGLNLLLSIVMLVDVTFGMALWDISVNAVSLINIVSAGVMSVRFVSHI 1185
Db 1111 VAVLLGCLNSAVINCATIAMILLVMFGVMMWGLISLNAVSLVNLWMSCGISVFECSHI 1170
QY 1186 TRPSAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVILGLAKAQIQLIFFRLNLL 1245
Db 1171 TRAFVTSTKGRVERAEALSHMGSSVFGSITLTKFGGIIVLAFAKSQLFIQFYFRWYLA 1230
QY 1246 IYLLGLLHGLVFLPVLSTVSGDDVNPALAL 1275
Db 1231 MYLLGATHGLIIFLPVLLSYIGPSINKAKSL 1260

RESULT 16
Q9TT75 RABIT
ID Q9TT75_RABIT PRELIMINARY; PRT; 1286 AA.
AC Q9TT75_2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Niemann-Pick type C1 disease protein.
GN Name=NPC1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
RN NCBI_TaxID=9986;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RA Murakawa M., Freeman M.W.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF202730; AAF18444.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:heghehog receptor activity; IEA.
DR GO; GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; I.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS50156; SSD; 1.
SQ SEQUENCE 1286 AA; 142723 MW; 9266CC310F64CCD7 CRC64;
Query Match 34.2%; Score 2361; DB 2; Length 1286;
Best Local Similarity 39.0%; Pred. No. 5.1e-158;
Matches 513; Conservative 237; Mismatches 455; Indels 110; Gaps 26;
QY 7 RGMILLWALLRLLAQSEPYTHIQPGVCAPFYDCGNKPNELSGSLMTLSNVSCLSNTPARKI 66
Db 13 RGRLLGLLALLLLCPAQLV-----PQSCWYGEQ---IASGD---KRYNCQYSGPKPL 60
QY 67 TGDHLILQIKCPRIYTGNTQACCSAKQLVSLASLSITKALLTRCPACSNFVNLHCH 126
Db 61 PKDGYDLIQELCPGFFP-DNVSLLCDVQQLQTLKDNLQLPLQFLSRCPSCFYNLMLFCE 119
QY 127 NTCSPNQSILFINTVTRVAQLGACQLPAVAY-----EAFQHFABQSYSDS 171
Db 120 LTCSPRQSQFLNVT-----ATEAYVDPATNQTNTYKELQYIIGORPADAMYNA 168
QY 172 CSRVRVPAATLAVGTMCGVYGSALCNQORWLFQDGTGNGLAPLDIT--FHLLEBPQAV 229
Db 169 CRDVEAPSSNDKALGLLCKGDANA-CNATNWIYMFDRKNGQAPFTITPISDLPV---- 223
QY 230 GSGIOPLENGVARCNESQGDVATCSQDCQCAASCPAIARPQ-----ALDSTFYLQ 280
Db 224 -RGMEPMNATKGCNEAVDEVTGCPSCQDCSVVCQPKQPPPPPIPRIFGLDAMYVIMW 282
QY 281 MPGSVLVLIILCSVAVVTILLVGRVAPARDKSKMWDPKGTSLSDKLSFSHT-- 335
Db 283 ITYMAFLPMFPGTFFAVMVCY-----RKRYFVSEYTPIDSNIAFVNTSDKGE 329
QY 336 -----LLGQFFQG-----WGTWVASWPLTILVSVI VALAAGLVFTELTTPDVEL 382
Db 330 ASCCDLLGAAPFGCLRLRFTRMGSCFVCRNPGCVIFPFGFIAACSSGLVYVVTNPIDL 389
QY 383 WSAFNSQARSEKAFHQHFGFPFRFTNOVILTPNRSYRYDLSLLGPK-NFSGILDLDL 441
Db 390 WSAFNSQARQDKKEFFDAHFGFPFRTEQLIIRAPHTNKHYQYPPSGADVPPGPDLDKQL 449
QY 442 LELLEQLERLHLQVMSPEAQRNLSQDICYAPLNPNTSLVDCCINSLLQYFQNNRTLL 501
Db 450 HQVLDLQTAIE--DITASYNNTVTLQDICTAPLSYPNK---NCTILSVLNTFQNSHML 504

```
QY 502 LLTANQTLMGQTSQVDMKDFLYCANAPLTKDGTALALSCMADYGAPVPPFLAIGYKG 561
Db 505 DHEQGDFF---VYADHTHEFLYCVRPASLNDTSLHDPCLGTGGFGGVPFWLVGLGYDD 561
QY 562 KYSEAEALIMTFSLNYPAGDPRLAQAQKWEAEFLEEMRAFQRMAGMFQVTTAERSL 621
Db 562 QYNNATATLVITFPVSNYNTEKQRAKAEKEFINPVKNYEN---PNLTISTATERSI 618
QY 622 EDEINRTTAEDLPATSYIVIFLYISALGYSYSGRVMVDSKATLGLGVAIVVLGAVM 681
Db 619 EDELDRENSDVFTVVISYAVWFLYISALGHKSCRFVDSKSLGIAGILVLSVA 678
QY 682 AAMGFYSYLRSSILVQVPPFLVSVGANIFIFVLEYQRLPRRGPPEVHIGRALG 741
Db 679 CSLGIFSYIGIPLTLIVIEVIFPLVAVGVNIFILVQTYQDERLOGETLDOQLGRVLG 738
QY 742 RVAPSMILCSLSEATCFPLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSLDSKR 801
Db 739 EVAPISIFLSSFAETVAFVFLGLSVNPAVHTFSLFAGMAVFDLLOITCFVSLGLDKR 798
QY 802 QASRLDVCCCKVQELPPGQ--GREGLLGFFQKAYAPPLLHWITRGVVLVLLFALFVGS 860
Db 799 QEKQLDVLCCPRGAEAGTSIQASESYLFRFPKNSYSPLLKDWMRPIVIAVFGVLFSFS 858
QY 861 LYSMCHIISVGLDQELALPKDSYLLDYFLFNRYFEVGAPVYVITLGYNFSEAGMNAIC 920
Db 859 IAVLNKVEIGLQSLMPDDSYVDYFKSLQGYLHAGPPVYVLEEGHNYTSLQGNMVC 918
QY 921 SSAGCNPFSTQIKIYATFEFQSYLATPASSWVDDFDLWLT--SSCRLYISGNKDKF 979
Db 919 GGLGCDNLSVQOIFNAALQNDYTRIGAPSSWIDYDFDWKQSSCCRV---SNVTQF 975
QY 980 CPSTVNSLNLCKNCSMIT--MGSVRFSVEQFHKYLFWLNDRENIKCPKGGAAYSTSVNL 1038
Db 976 CNASVVDPAVCR--CRPLTEGQRQGEDFMFLPWLSDNPNKCGKGGHAAVGTAVNI 1034
QY 1039 TSDG--QVLSRPMAYHKPLKNSQDYTEALRAELAANITADLRKVPCTDPAFEPFYTI 1097
Db 1035 LGNGTEVGATYFTVHTVLOTSAFDIDAMKKAQLIAGNATATM--GLKGS--RCRVFPFYSV 1091
QY 1098 TNVFEQYITLPEGLFMLSCLVPTFAVSCLLGLDLRSLGILNLLISVMILVDTVGEMA 1157
Db 1092 FYVFEQYITLDDAVNFGVSLGAVFVAVVAVVLLGCELWSAVIMCVTITMILVNMFGVM 1151
QY 1158 LMDISYNAVSLNLSVAGSVSEFVSHITRSFAISTKPTWLBRAKATISMGSAVPAGA 1217
Db 1152 LMGISLNAVSLVNLVMSCGISVEFCSHITRAFTVSGKSRVARAEALAHMGSSVFSGIT 1211
QY 1218 MTNLPGLVGLAKAQLOIEFPELNLITLLGLLGLVFLPVLSYVGPVNP 1272
Db 1212 LTKFEGIVVLAFAKSQIFQIFPYFRMYLAMVLLGATHGLIFLPVLSYIIGPSTNKA 1266
```

RESULT 17

```
Q4RW5_TETNG PRELIMINARY; PRT; 1209 AA.
ID Q4RW5_TETNG PRELIMINARY;
AC Q4RW5;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Chromosome 15 SCAF14981, whole genome shotgun sequence.
DE (Fragment).
DE ORFNames=GSTENG0027641001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
```

```
RA Mauceli B., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Kattinka M., Vacherie B.,
RA Blument C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kallis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014981; CAG07097.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 1209
FT SEQUENCE 1209 AA; 133913 MW; 2046F9D3DE4E48F1 CRC64;
SQ
Query Match 33.8%; Score 2338; DB 2; Length 1209;
Best Local Similarity 41.5%; Pred. No. 2e-156;
Matches 520; Conservative 217; Mismatches 414; Indels 102; Gaps 30;
QY 72 ILLOKICPRLYTGPTGNTQACCAKQVLSLEASLSTKALLTRCPACSNFVNLHCHTCTSP 131
Db 1 VCLQELCPG-YDYGNSRLCCDVNLHTLKEVLPQLFQSRCPACFFNLNLPCLCTCSP 59
QY 132 NQSLPIWTRVRAQAGQLPAVAYEAFYQHSFAEQSDSCSRVRAAATLAVGTWGV 191
Db 60 HQSQFMATKLSG-----PDWVEQYIYIGLTFANMTNACRDVQAESSNVKALSILCG- 112
QY 192 YGSALCNAQWLNFGQDTGNGLAPLDTFHLLEP--QOAVSGSIGIPLNEGVARNESQGD 249
Db 113 KDAKCHATNWIQWNTDNEQAPFIT-----PIPSDVPVSGVTPMNDTAACTDGLD 167
QY 250 DVATCSCQDCAASC-PAIARPAQALDSTFYLGQMGSLVLIILCSVFAVTVILLGVFA 308
Db 168 GSGFCSCQDCTNACGPRVPPPTPAWKILGM---DAMTVIMWFSYMAFLIFVGSLLIA 224
QY 309 PARDK-----SKMVDPKKCTSLS---DKLSF-----STHTLLGQFFQG----- 343
Db 225 WCHRKETSIMYGPFLDSKNRPSLNRPNDPDPVFPFLDDASCCTELSERFESYLRSCFSC 284
QY 344 WGTWVASWPLTILVLSVIPVVALAAGLVFTBELTDPVELWSAPNSQARSEKAFHQHFGP 403
Db 285 WGSFCVLPVCLVLLGSLILVASSGGLIYMRITDPDVLWSPPSSQARQERYFDSHFPG 344
QY 404 PFRNQVILTAPNRSYRYDSLLGPK-NFSGILDLDLLELLELLELLELLELLELLELLE 462
Db 345 PFRTAQLIITSPLNDTFLYSPVMGPDIPFKAVLSKDLHQVLDLQLDIESL--VATYEQ 403
QY 463 RNISLQDICAPLNPNDTSLVDCCLNSLQYFQNNRTLLLTANOTLMGQ-TSQVDWFKDH 521
Db 404 -SVTLKDICLAPLSYND---NCTILSVLNFQNSHS-----TLDHVLKDFLVYADFHS 455
QY 522 FLYCANAPLITFKDGTALALSCMADYGAPVPPFLAIGYKGDYSBAELIMTFSL-NNYP 580
Db 456 FLYCVSAPASLNDTTPHDPCLGTGGVPFWLALGGYDDTNNNATLAVVTFPINNNY- 514
QY 581 AGDP-RLAQAQKWEAEFLEEMRAFQRMAGMFQVTTAERSLEDEINRTTASDLPIFATS 639
Db 515 --DPTKGLKTLAWEKEFIRFMKNYSN---FNUTIASAERSIEDEINRESNDISTIVVS 569
QY 640 YIVIFLYISLALGYSYSSWSR-----VNVDSKATLGLGVAVVVLGAVMAAMGFY 689
Db 570 YVIMFVYISLALGHQTSFRLPLPHVLLLLLLLVDSKVSIGILVLSVSSSLGIFSY 629
```


DR FlyBase; FBgn0024320; NPC1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008159; F:hedgehog receptor activity; IEA.
DR GO; GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1287 AA; 143290 MW; 540AAB292A37790C CRC64;

Query Match 32.7%; Score 2262; DB 2; Length 1287;
Best Local Similarity 37.8%; Pred. No. 5.3e-151;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LLWALLRLAQSEPTTIHQPGYCAFYDECGNPELSGLMTLSNVSCLSNTPARKITGD 69
DB 17 LIAAVLFTLIQSKD-----CVYGVCTN-DPSHS-----QNCYPNGTAKEMATD 62

QY 70 HLILQKICPRLYTGNTQACSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNWC 129
DB 63 GLELLKRCGFLLENSKFCCKNQVELLNKNVELAGNILDRCPCSMENLVRHICQPTC 122

QY 130 SPNQSLEFINTVRAQAGQLPAVVAEYAFYOHSAEQSYDSCSRVRPAAATLAVGTMC 189
DB 123 SPKQAEFMHVATQKNKKGD-BYISGVDLHIHSTYINKTYKSCQSVQPGQLAFDLMC 181

QY 190 GYVGSALCNAORLWFOGDTGNGLAPLDITPHLEPGQAVSGGIQPLNEGVARCNESQGD 249
DB 182 GAYSARCNPTKWFPMGDATPYVPFQIITYIQHEP-KSNSNNFTPLNVTTPCNOAVSS 240

QY 250 DVATSCDPCAAASC-----PAIARPQ-----ALDSTFYLGQMPGSLVLIILCSVPAV-VT 299
DB 241 KLPACSCSDLSLCPQGPPEPRPPEPKIVGLDAVF-----VMAAFLVLVGL 288

QY 300 ILLVG---FRVAPARDKSNVDPKGTSLSKLSFS-----THLLGQFFQGW 344
DB 289 VFLMGSLFTQGSMDNDNFQVD---GNDVSDMPYSENDSYFEKLGATHTETLETFYTKW 345

QY 345 GTWVAGSPLTILVSVIPVVAALAGLVFTELTTDPVLEWSPANSQARSEKAFDHQFGPF 404
DB 346 GTYFASNPGTLTLAGASLVILGYGINFIEITDPVKLWASFNKSRLEREFDTKFSFP 405

QY 405 FRTNQVILTAPNRSSVRYDLSLLGPKNFSGIIDLLELLELQERLHLQWSPQAQRN 464
DB 406 YRLEQIIKAVNLPQIVHNT-SNGPYTFGVPFDRFLTKVLDLQEGIKEINA-----NG 458

QY 465 ISLQDICYAPLNPDTSL--YDCCINSLLQYQNNRTLLLTANQTLMGQTQSDVDKDH 522
DB 459 TQLKIDICYAPLSDGSEIDVSCVQWSIGYFGDDRRL-----DDHDENGFNVTYLDAL 514

QY 523 LYCANAPLTPKDGTLALSCMDYDGYAPVPFLAIGY-----KGKDYSEAEALIMTF 574
DB 515 YDCISNPLY-----CLAYGPGVPDPAIALGGLFPDGLTGTSTKPELANAIIITF 564

QY 575 SLNYPAGDPRLAQAKLWEAFLEEMRAP-ORRMAGMFQVTFTAERSLSEDEINRTAEDL 633
DB 565 LVKNH-HNKTDLLENALTWEKKFVEFMVTNTYKNNMSQYMDIAFTSERSIEDLNRESQSDV 623

QY 634 PIFATSVIVFLYISLGLSYSSWSRWMDSKATGLGVAVLGAVMAAMGPFYVLGR 693
DB 624 LTLVSVLYIMFMXIALSHGVKFKRVFIDSKITLGLGGVIVLASVSVSGVFGYIGLP 683

QY 694 SSLVILQVFPVLVSGADNIFIVLEYQRLPRRPGEPREVHTRALGRVAPSMMLCSUS 753
DB 684 ATLIIIVEIFVLVAVGVDFNIFILVQTHQRDKPNETLEQQVGRILGKVPSSMLTSL 743

QY 754 EATCFFLGALTTPPAVPTFALTGSLVILDFILQMSAFVALLSILDSKROEASLQVCCCV 813
DB 744 ESFCFFLGLSDMPAVRAFALYAGVALIIDFLDITCIVSLFTLTKRREENRMDICCFI 803

QY 814 KPOELPPPGQEGLLGLFFQKAYAPFLHLHWITRGVVLLFLALFGVSLXSMCHISVGLDQ 873
DB 804 KGKKPDSITNEGLLYKFFSSVTVFPMKKIVRASVMVFFAWLFCFIALAPRIDIGLDQ 863

QY 874 ELALPKDSYLLDPLFLNRYFFEVGAPVYFVTTLYGYNFSSBAGNAICSSAGCNFSFTQK 933
DB 864 ELAMPQDSFVLHYFQSLNENLNIGPPVYFVLKGLAYTNSDQNLVCAGQYCNDDSVLTQ 923

QY 934 IQYATEPPEQSALAPASSWVDDFDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNL 990
DB 924 IYLAHSHSNQTYIARPASSWIDDYFDMAAASCCCKY-----RKDSGDFCPH--QDTSL 976

QY 991 KNCMSITMGS-VRPSVEQFHLYFWFLNDRPNIKCPKGGLAAYSTSVNLTSDDQVL--A 1046
DB 977 R--CNITKNSLLARPEEKFKVLYLPFFLKNDPDDTCAKAGHAAYGAVRYSNSHERLNEA 1034

QY 1047 SRFMAYHKPKNSQDYTEALRAARELANATADLR-----KVP-GTDPAPFEPYPTITN 1099
DB 1035 SYFMAYHTILKSSADYFLALESARKISANITQMLQGRLMNGVPMASALTVEVPFYSVFY 1094

QY 1100 VFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLLNLLSIWMLVDTVGFMAJW 1159
DB 1095 VFYEQVLTWMSDTHLOSNGISVLISFVTVFVLMGFVHSALVVVITITWIVNVLGGLMYW 1154

QY 1160 DISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMT 1219
DB 1155 NISLNAVSLVNLVAVGISVEFCSHLVHSPATSKSVSQIDRAADSLSKMGSSIFSGITLT 1214

QY 1220 NLPQILVLGLAKAQLIQIFFRNLNLLITLGLLGLVFLPVILSVYVGPVDVNPALALBQKR 1279
DB 1215 KFAGILVLAFAQSIFQVYFPRMYLGVIGVIAHGLIFLPVLLSYIYGAPVSNARLYHSQ 1274

QY 1280 A---BEAVALVM 1288
DB 1275 AAETHETALAGIL 1287

RESULT 19

QYU5W1 DROME PRELIMINARY; PRT; 1287 AA.
AC QYU5W1; DROME PRELIMINARY; PRT; 1287 AA.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NPC1 protein.
GN Name=NPC1; ORFNames=CG5722;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang Y.Q., Xue H., Broadie K.S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249606; CAB56505.1; -; mRNA.
DR FlyBase; FBgn0024320; CG5722.
DR FlyBase; FBgn0024320; NPC1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008159; F:hedgehog receptor activity; IEA.
DR GO; GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1287 AA; 143277 MW; 8E154FE5E5DEF556 CRC64;

Query Match

32.7%; Score 2256; DB 2; Length 1287;

Db	63	GLELLKRCGFLLESENKFCDDKNQVELLNKNVELAGNAILORCPSMENLVRHICQPTC	122
Qy	130	SPNQSIFINVRVAQIAGAGOLPAVAYAEAFYQHSFABQSDYSCSRVRUPAAATLAVGTMC	189
Db	123	SPKQABFMHVATQKNKGD-EYISSVDLHISTEYINKTYKSCQSVPQTQQLADFLMC	181
Qy	190	GUVYSALCNAQRWLNFQDGTNGGLAPLDITLHLPQQAAGVSGIOPLNEGVARCNEQCD	249
Db	182	GAYSAGRCNPTKWNFMGDATNPYVPQIYIQHEP-KSNSNNFPLNVTVPQCNQAVSS	240
Qy	250	DVATCSQDCAASC---PAIARPQ-----ALDSTFYLQMGFGLVLIILCSFVAV-VT	299
Db	241	KLPAACSCDCLSCPQCPPEPPPEPKVGLDAYF-----VIMAVALVGLV	288
Qy	300	ILLVG---FRVAPAROKSKWDPKKGTSISDKLSFS-----THLLGQFFQCM	344
Db	289	VFLMGSLFTQGSMDNFQVD---GNDVSDDEMPYENDSYPEKLGHAHTETLETFTFKM	345
Qy	345	GTWVASPLTILVLSVTPVVALAAGLVFTLTPDVELMSAPNSOARSEKAFHDHFGPF	404
Db	346	GTYFASHPGLTLAGASLVILGVGNFIETTPVKLWASPNKSLREBREFDIKFESF	405
Qy	405	FRTNQVILTAPNRSSYRSDLLGPKNFSGILDLDLLELLELLELQERLRLHQWSPQAQRN	464
Db	406	YRLBQIIIKAVNLPIQVHNT-SNGPYTFGPVDFREFLTUKVLDLQEGIKEINA-----NG	458
Qy	465	ISLODICVAPLNPDNTSL--YDCINSLLQYFQNNRTLLLTANQTLMGTSQVDWKDHF	522
Db	459	TQLXDICVAPLSDGSDIDVSCQVQWQIMGYFGDDRERL-----DDHEDNGFNVTYLDAL	514
Qy	523	LYCANAPLTFKDGITALSADWAGVPEFLAIGVY-----KGKQYSEAEALIMTF	574
Db	515	YDCISNPYL-----CLAPYGVDPDAIALLGFLPEQDQLTGSTKPELANAILTF	564
Qy	575	SLNNYPAGDPRLAQKLWEBAFLREMAF-QRRMAGMFQVTFTAERSLEDEINRTTAEDL	633
Db	565	LVKXH-HNKTDLLENALTWEKFEVFETNYTKNNMSQYMDIAFTSERSIEDELNRESQSDV	623
Qy	634	PIFATSYIVPIXYTSLAIGSVSSMSRWSDSKATLGLGVAVVLGAVMAWGMGFFSYLGIR	693
Db	624	LTILSVYLIMFYIAISLGHVKFEPFVIDSKITLGTGGVIVILASVSVSGVFGYIGLP	683
Qy	694	SSLVTLQVVPVLVLSVGDADNIFIVLEVRQLPRRPGPREVHIGRALGRVAPSMLLCSLS	753
Db	684	ATLIIVEVPEFLVAVGVDNIFILVQTHQRDQRPNETLEQQVGRILGKVGPSILLTSL	743
Qy	754	EACFFFLGALTPMPAVRTFALTGSLVILDFLLQMSAFVALLSIDSXQSEASRLDVCCCV	813
Db	744	BSFCFFGLGLSDMPAVRAFALYAGVALLIIDPLLQITCFVSLFTLDTKRRENRMDICCFI	803
Qy	814	KPOELPPQCGEGILLGFQKAYAPFLHWTIRGVVLLLLFLALPGVLSYMSCHISVGLDQ	873
Db	804	KGKKPDSITSNEGLLYKFFSSVYPFLMKKTVRASVMVIFPWLJCFISIAIPRIDIGLDQ	863
Qy	874	ELALPKDSVLLDYFLFLNRYFEVGAUVVFTTGLGNFSSSEAGMNAICSSAGCINNFSQK	933
Db	864	ELAMPQDSFVLHYFOSLDENLNGPPVYFVLKGLDAYTNSDDQLVCAGQYCNDDSVLTQ	923
Qy	934	IQYATFEFQSYLAIPASSWDDDFIDW-LTPSSCCRLYISGPNKDK--FCBSTVNSLNC	990
Db	924	IYLASRHSNQTYIARPASSWDDYFDMAAAASSCKY-----RKDSGDFCPH--QDTSCL	976
Qy	991	KNCMSITWGS-VRPSVEQHKYLPWFINDRPNIKCPKGLAAYSTSVNLISDGOVL--A	1046
Db	977	R--CNITKNSLRBEKFKVLYPEFLKNDPDDTCAGAHAAYGAVRYSNSHERLNEA	1034
Qy	1047	SRFMYHKPLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAPEVEFPYITN	1099
Db	1035	SYFMAYHTLLKSADYFLAESARKISANITQMLQORLMSGVPMASALTEVEFPYSVFY	1094
Qy	1100	VFEYEQYITLPEGLFMLSICLVPYTPAVSCLLLGLDLRSLGLNLLISVIMLVDTVGFMAW	1159

1095 VFYEQYLTKWSDTLQSGNISVLISFVTVFLMGDFVHSALVVVITITMIVVNLGGLMYW 1155

1160 DIGNAVSLINLVSAVGMSEVFVSHITRSPALSTKPTWLERAKEATISMGSAVFAGVMT 1219

1155 NISLNAVSLVNLVWAGVISVFEPCSHLVHSPATSKSVSQIDRAADSLSKMGSSIFSGITLT 1214

1220 NLPGILVLGLAKAQLIQIPFFRLNLLITLLGLLHGLVFLPVILSYGPDVNPALALEQKR 1279

1215 KPAIGILVLPAPKSIQFQVFFRMYLGIWVIGAAGHLIFLPLVLLSYTGAPVSNARLRYHSQ 1274

1280 A-----EAAVAVM 1288

1275 AAAEHETALAGIL 1287

RESULT 21

Q7Q409 ANOGA

ID Q7Q409 ANOGA PRELIMINARY; PRT; 1291 AA.

AC Q7Q409;

DT 01-MAR-2004 (TrEMBLrel. 26, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ENSANGP0000010415 (Fragment).

GN ORFNames=ENSANG00000007926;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Culicidae;

OC Anophelinae; Anopheles.

OX NCBI_TaxID=180454;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT "Anopheles gambiae re-annotation.";

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AAAB01008964; EAA12360.2; -; Genomic DNA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008158; F:hedgehog receptor activity; IEA.

DR GO; GO:0030301; P:cholesterol transport; IEA.

DR InterPro; IPR004765; NP_C_type.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR000731; SSD_5TM.

DR Pfam; PF02460; Patched; I.

DR TIGRFAMs; TIGR00917; 2A060601; 1.

DR PROSITE; PS50156; SSD; 1.

FT NON_TER

FT SEQUENCE 1291 AA; 142230 MW; 244A70BEDDA3F519 CRC64;

Query Match 31.9%; Score 2205; DB 2; Length 1291;

Best Local Similarity 37.4%; Pred No. 5,9e-147;

Matches 491; Conservative 242; Mismatches 470; Indels 110; Gaps 31

QY 33 CAFYDECGKNPELSGLMTLSNVSCLSNTPARKITGDLHLLIQKICPRLY-----TGPNTQ 88

DB 8 CVWYGIC--NVDSLGRAQ-----YCPYNGTKAIDTEASDLLKWKCKHLLVEDATG-SIN 59

QY 89 ACCSAKQLVSLASISITKALITRCPACSDNPNVLHCHNTCSPNOSLFINTVRV----- 142

DB 60 TCCDAAQVKVLNKNVALAANFUSRCFPCSMANLVRMHCDFTCSFQSSFMKVVSTBEVDQ 119

QY 143 -----AOLGAGOLPA--VVAYEAFYQHSFAEQSYDSCSRVRYPAATLAVGTMCVY 192

DB 120 EGESEKADATTAAASPKEYIITKIDHITQQYLNGTFBSCNQSVPSTCOLALDLMCGDW 179


```
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ENSANGP0000007352 (Fragment).
GN ORFNames=ENSANG0000005547;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008846; EAA06340.3; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgheg receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; P550156; SSD; 1.
FT NON_TER 1
FT NON_TER 1003
FT NON_TER 1003
SQ SEQUENCE 1003 AA; 111208 MW; P6B74BD2054D9E9A CRC64;

Query Match 25.5%; Score 1760; DB 2; Length 1003;
Best Local Similarity 38.3%; Pred. No. 1.5e-115;
Matches 402; Conservative 190; Mismatches 385; Indels 72; Gaps 23;

QY 243 CNESQDDVATCSCODCAAPARPOALDSTFYLQMPG---SLVLIILCSYPAVVT 299
DB 1 CNEY-DGYSACVDCDESEPTSEPPQPKDPGFVMVDLNGVTFVAVVVGIGLACIVL 59
QY 300 ILLVGRVAPARDKSRVDPKKGTSLSDKLSP-----STHTLLGQFPQGWGTWASWPLT 354
DB 60 ALLFG-----GREGSKQ-----QLPDLPSFGGPPSVNRALGRFTWGTGFCARNPVL 107
QY 355 IIVLSVIPVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILPA 414
DB 108 ILAICSWIVGGLAFGIQYLIITTDVPELWAPDSRARQKDYFDSRFSFPRTEQIFI-K 166
QY 415 PNRSSVRYDSSLGPKNFGSILDLILLLELLOERLHLQWSPEAQRNLSLQDICYAP 474
DB 167 PTRQEFHHTAGNQTPGAVDREFLEVFVKLTQTTIE--QLGQAEGR---GLEQICFAP 221
QY 475 LNP--DNTSLYDCINSLILQYFQNNRLLILLTANOT---LMGQTSQVDMKDFLYCANAP 529
DB 222 MTQVGAQTVLSECTQSVFGYFENS---LAFENRTGDLNGFV--VNYLDKINGCTR-- 273
QY 530 LTFKQGTALALSCMADYGAPVFPFLAIGY-----KGHDYSEAEALIMTFSLNNYPAGD 583
DB 274 -----NAYLPSCFGTYGGPIEPGAVGGFPQAPGANDPRLATGVLTFLVEN-KANK 326
QY 584 PRLAQKLEEAFLERAFQRMAGMFQVTTAERSLEDEINRTTAEDLPATSIYVI 643
DB 327 DELGPALEWEQRFVDMRDFEH---PMMEVAVSAERSIEDGDEMSAEAMYTVIISYVM 383
QY 644 FLYISLALSGYSWSRVMDSKATLGLGVAVVLGAVMAAGFFSYGLRSSLILQVVP 703
DB 384 FVITFSLKINGFRQFLHGRIVLAVGGIVVVLASVACSLGFFGYGLEATMTMLTIEVP 443
QY 704 FLVLSVGDADNIFVLEVQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEACFFLGL 763
```

```
DB 444 FLVLAVGVNDNIFMLVHAFNRIDRERTPETACATGEALGQISILTSSECCFAIGAL 503
QY 764 TMPAVRTEALTSGLAVIDLFLQMSAFVALLSLDSKROBASRLDVCCCV---KPOBLPP 820
DB 504 SMPAVNTFANTATVALLVFLQISAFVALMDEQRVERGLDQVCCVRATNKPKPV 563
QY 821 PQQEGLLGFPQKAYAPFLHMTIRGVVLLFLALFGVLSYSMCHMCHISVGLDQELALPKD 880
DB 564 APDQPGWLERVVERFVVPFLPRVLTVALFLVWGSLSLMVVPSEFGLDQELSMARD 623
QY 881 SYLLDYFLFNRYFVGAPVYFVTTLGYNFSEAGNAICSSAGCNFQTKIQVATEF 940
DB 624 SHVVKYFPEMAELFMWGPVYFVVKAGLNYTDVQHQLVCGGLTCLNDDSIKLYQASLY 683
QY 941 PQOSYLAI PASSWVDDFDLWLPSSCCRLYISGPNKDKPCPTVNSLNCNKMCSITMGS 1000
DB 684 PETHIARPASSWLDYIDWLAIQSCCKY---NPTDGSFCAS--NIVFCAPCEEYDETG 738
QY 1001 VRPSVEQPHKYLFWFLNDRPNIKCPKGGLAAYSTSVN--LTSQDQ--VLASRFMAYHKPL 1056
DB 739 IRPTVPQPERYLEFFFLSDLPDENCAKAGRAAYSRAALNYLLDRGLNVQDSYFSTVHTTA 798
QY 1057 KNSQDYTEALRAARELANITADLRKVPCTDPAFVFPPTITNFFVEQYLTILPEGLFML 1116
DB 799 VTSRQFYTALEQARLIAADIQOMLDE---RQAGVEIFPYSVFVFEQYLTIVSDALQSL 855
QY 1117 SLCLVPTTAVSCLLGLDLRSGLNLSIVMTLVDTVGFMAWDISYNAVSLNLSVAVG 1176
DB 856 GLSLAAVFTVTLVGLDLSLVLLVFLVLMNMLNLNITLNAISLVNLVMSVG 915
QY 1177 MSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQ 1236
DB 916 IGVFEFISHIVRTYRLA-HGTRLESAAMVTRTSSSVFSGITLTKPAGIIVLAPAKSQIFQ 974
QY 1237 IPFFRLNLLITLGLLHGLVFLPVLSYV 1265
DB 975 IFYFRMYCLIVLVAAGHLILLPVFLTYI 1003

RESULT 24
QSLNK7 MAGGR
ID Q51NK7 MAGGR PRELIMINARY; PRT; 1275 AA.
AC Q51NK7;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG00819.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birten B., Nussbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blichsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheehatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Horan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamat M., Kamat A., Kamysseis M., Karlsson E.,
RA Kellis C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
```


RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinska M.I.,
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Perteau M., Riggs P.R., Salzberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
Cryptococcus neoformans";
RL Science 307:1321-1324 (2005).
DR EMBL; AB017345; AAW43643.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TH.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1330 AA; 146237 MW; 223B9F89F35EB5A5 CRC64;

Query Match 24.0%; Score 1659.5; DB 2; Length 1330;
Best Local Similarity 30.6%; Pred. No. 2.9e-108;
Matches 427; Conservative 247; Mismatches 448; Indels 275; Gaps 40;

QY 31 GYCAFYDEGKPELGGSLMTLSNVSLCNTARKITGDHLLIKICPRLYTGNTQAC 90
DB 11 GICMRGTGK-----TSMGADLPDSDSDATPDQRLDLMAVSGPSYSLPD-HVC 63

QY 91 CSAKQLVLEASLSITKALLTRCPACSNFVNLHCHNTCSNQSLFINVTRVAQAGQOL 150
DB 64 CTYDQLSTLSDRLLQAAPLIASCPCINFRSFCYDFTCSNQSLFSLVTATQKTEGK- 122

QY 151 PAVAYEAFYQHSFPAQSDSCSRVRVPAATLAVGTGCVGSGALCNAQRLNFGQDTG 210
DB 123 DAVKEVDYEVSSDFKGFYDCKQVFGATNGFAM-----DLTGGCATNASGFLKTMGDLR 178

QY 211 NGL-APLDTITLHLLPQAVSGIOLNEGVA-----RCNESQDDVATCSCQCAASC 263
DB 179 PGLGSPFQINP-----PDNDDSYARRAPLSCSDAENIN-ARCAACDPCSV 223

QY 264 PA---IARPOALDSTFYLGQMPGSLVLIILCSVFVAVTILLVGRVAPARDKSK----- 315
DB 224 PSLPIAIPST--KQCHVGAVSCLTFLSIITYSVIILGALLYIWKQA-AHRQRBYERV 280

QY 316 -MVDPK-----KGTSLSDKLSF-----S 332
DB 281 ALLDDPHSPFTIQNGGSLDGLMGRNDDAESGPSGIHFRLCRGASLDDPMHLQPKQNK 340

QY 333 THTLGGQFGQGTWVWASWPLTILVSVIPVALAAGLVFTLTDPVELWSAPNSQARS 392
DB 341 INATLRRFFYRLGLTCAKREVEFAISALI VGLFNFGWKYFEVETDPVRLMVSPSESAS 400

QY 393 EKAFHDFHGFPPFRNTNQVILTAPNRSSVRYDSLLIGPKNFSGILDLLELLELLEQLRL 452
DB 401 QKQFPDSDFGFYKSEQIFITQSSGSPVNYDT-----LDWLVK----- 439

QY 453 HLQWNSPEAQRNISQDICYAPLNPDNSTLYDCCINSLLQVFNQNNRLLTLLTANQTLMGQ 512
DB 440 ETEINALKTSQDGLIEDICFAPAGKT---PCVIOVSVAW-----LGD 479

QY 513 TSQV---PWKHFLYCANAPLTFKDGATALALSCMADYCAPVFPFLAIGYKGYKDYSEABA 569
DB 480 DMEVNGEWESESRVDCASRP-----GECLEPPFGQPIDPKLVLGANG-DWLEAKA 528

QY 570 LIMTSLNNYPAGDPRLAQAQKLWEBAFLTEEMRAFORMAGMQFVTFTAERSLEDEINRTT 629
DB 529 LVVTVWVNNY--NDERVEPAEQWERKLDYLRDLKR--PGI-KISVSTGVSLSEEBINKST 583

QY 630 AEDLEIPATSYIVIFLYISLALG-----SYSSWSRV----- 660
DB 584 NTDVKIVLVSVMFFYSYLSLTGGLPPSMIOQAFTHRAYRLVFRIGLVLLHVKDAPLES 643

QY 661 -----MVDSKATIGLIGGAVVVLGAVMAAMGFSSVGLGRSSSLVILQVVPFLVL 707
DB 644 APPDPFRVIPITLLSVNSKESGLFGIVIVLIVAVSSVGLFSLGLVGRVTLIIAEVIFPVL 703

QY 708 SVGADNIFIFVLEYOR-----LPRPGBPREVH-----IGRALGRVAPS 746
DB 704 AVGVDNVFVILVHELERQNNLHAAQCPGDDSVHSGAQPSTGTFLAPEERVARAVARMGFS 763

QY 747 MLLCSLSAICFFICALTPMPAVRTFALTSLGLAVLDFLLOWSAFVALLSLDSKQESAR 806
DB 764 IMLSSVTEVAFALGALVMPAVRNFAIYAGSVLFGAIMQCTVFVSAMTLDLRESMR 823

QY 807 LDVCCVKPQBLPPPG-----QCEGLLGFQKAYAPFLHMTTRGVVLLLELALFCV 859
DB 824 IDCPCIRLR--PPGLYDNEAPREGWVKFMRVTYAPSLRHEVKQLVAVFAGLFLA 881

QY 860 SLYSMCHISVGLDQELALPKDSYLDYFLFLNRYFEVGPVYFVTLTYNFSESEAGMNAI 919
DB 882 AIIQIHSISGLDQRLALPSSHLVPYFNDVDSYLDVGPVYFVTE-GGDPSSRHQQQL 940

QY 920 CSS-AGCNPFSTQKIQYATEFPPEQSYLAIPASSWDDPIDWLTPS--SCCR-----L 969
DB 941 CGREFTTCLLSVANTLEARKRPDSSFTSAPPAWIDDLOWNTNPTFSSCCRVRDRPTI 1000

QY 970 YISGNKDKFC-----PSTVNSLNCNCHSITWGSVRSVPEQFHKLPWFLNDRP 1020
DB 1001 FCSPKDSERLCRCPCEGQKWDSTMDGL-----PEGEDFMYLDEQWLISPT 1045

QY 1021 NIKCPKGLAAAYSTSVNL-TSDGOVLASRFMAYHKPKNSQDYETALRAARLAANITAD 1079
DB 1046 NDECPGQQAQPYSAVKLVSSNTTVAASHFRTYHTPLASQADFINALAAARRISDIT-- 1103

QY 1080 LRKPGTDPAPVEVPPYTTINPYEQYLTILEGLPMLSLCLVPTPAVSCLLGLDLRSL 1139
DB 1104 -----HRTGVRVPYSLFYFFDQYEHIIISMAIQVFLAFVAVLVTITSLIG-SWRITGG 1156

QY 1140 LNLISIVMLVDTVGPMALWDISYNAVSLINLVSAGMSVRFVSHITSP----- 1189
DB 1157 TVTFTCALAVINVMGVMGYWGSLNLSLNLVLSLGIATVBFCSHARAFMGAGSLPLD 1216

QY 1190 AISTKPTWLERAKATISMGSAVFAVAVTANLPGILVLGLAKAQLIQIPFRLNLLITLL 1249
DB 1217 KVEGRKEDERAWTALVDVGPVSFGITWTKLIGISVLALTSKLELVYFPMWLSLILS 1276

QY 1250 GLHGLVLPVLISVVG 1266
DB 1277 GALTGLVLLPVLISVYG 1293

RESULT 28

Q55SD4 CRYNE

ID Q55SD4 CRYNE PRELIMINARY; PRT; 1330 AA.

AC Q55SD4;

DT 13-SEP-2005 (TremBLrel. 31, Created)

DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TremBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=CNB2480;

OS Cryptococcus neoformans var. neoformans B-3501A.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;

OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.

OX NCBI_TaxID=283643;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing";
RL Submitted (JUL-2004) to the ENBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC ENBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01000024; EAL20887.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1330 AA; 146208 MW; 77DD1023070B520A CRC64;

Query Match 24.0%; Score 1658.5; DB 2; Length 1330;
Best Local Similarity 30.6%; Pred. No. 3.5e-108;
Matches 427; Conservative 246; Mismatches 449; Indels 275; Gaps 40;

QY 31 GYCAPYDCGKNPELSGSLMTLSNVCLSNTPARKITGDHLLQKICPLRYTGPNTQAC 90
DB 11 GICAMRGTCGK-----TSMFGADLPCDDSDATVPDQSLDLDMSVCGPSYSLPD-HVC 63

QY 91 CSAKQVLSLEASLSITKALLTRCPACSDNFVNLHCHNTCSNQSLFINVTRVAQAGAGQL 150
DB 64 CTYDQLSTLSDRLQAAPLIASCPACINFRSFYCDFTCSNQSTLSVTATQKTTEGK- 122

QY 151 PAVAYEAFYQHSFAEQSDSCSRVRVPAATLAVTGTCGVGSLCAQWALNFQDGTG 210
DB 123 DAVKEVDYEVSSDFKQGFYDCKQVQFATNGFAM-----DLGGGATNASGLKTYMGDLR 178

QY 211 NGL-APLDTTHLLPFGQAGVGGIOPLNEGVA-----RCNESQDDVATCSCDQCAASC 263
DB 179 PGLSGFFQINF-----PDNDDSYRRAPLSCSDAENIN-ARCAADCPSCVC 223

QY 264 PA---IARQALDSTFYLGQPGSLVLIILCSVFAVTVLLVGRVAPADKSK----- 315
DB 224 PSLPIAEPST--KQCHGAVSCLTSLIIYSVILVYKQALYIWKQA-ARHRQRYERV 280

QY 316 -MVDPK-----KCTSLSDKLSF-----S 332
DB 281 ALLDDPHSPSTIQNGQSGLDGLGRNDDAEGSPGSIHFLRGASLDDPMEHLQPKNK 340

QY 333 THTLGQFFQGGTGWASWPLTILVLSVPVVALAAGLVFTLTTPDVELMSAPNSQARS 392
DB 341 INATLRRFRRYRLGLTCAKPIEVFAISALVGLFNGWKYFEVETDPRVLWVSPSESAS 400

QY 393 EKAFHDQHPFPRTNQVILTAAPNRSSRYSDSLLLGPKNFSGILDLLELLELLEQLRL 452
DB 401 QKQFFDSTFGPFYKSEQIFITQSSGSPVNYDT-----LDWMLKV----- 439

QY 453 HLQVMSPEAQNRNLSLODICVAPLNPNTSLYDCINSLAQFQNNRTLLLTANQTLMGQ 512
DB 440 ETEINALKTSDDGLGLEDICFAPAGKGT---PCVIGSVSAW-----LGD 479

QY 513 TSQV---DWKDHPLYCANAPLTPKDGTALALSQADYAGAPVFPFLAIGGYKGYSEABA 569
DB 480 DMEVGEESRVRDCAARP-----GECLPPFGQPIDPKVLGGANG-DWLEAKA 528

QY 570 LIMTFSLNYPAGDPRIQAOKLWEBAFLBEMRAFQRMAGMFGVTTAERSLEDEINRT 629
DB 529 LVVTVWVNNY--NDERVEPAQWERKLDYRLDKR--FCI-KISYSTGVSLEEEINKST 583

QY 630 AEDLPATSVIVFLYVLSIALG-----SYSSWSRV----- 660
DB 584 NTDVKIVLSYLVNMFVVSUTLGGGLPPSMIQAFTHRAVRLVFRIGVLLHLVKDAPLES 643

QY 661 -----MVDSKATILGGVAVVLGVAAMAAGFFSYLGIRSSVLIVQVFLVL 707
DB 644 APPDFRVPITLLSVNSKPSLGLFGIVIVLIAVSSSVGLFSLIGVRLIIAEVFLVL 703

QY 708 SVGADNIFIFVLEQV-----LPRPGEPREVH-----IGRALGHVAPS 746
DB 704 AVGVNDVFIILVHELRQNNLHAAQQPGDDSVHSGNAQPSGTTFLAEEVARAVARWGS 763

QY 747 MLLCSLSEATCFFLGALTPAVRTFALTSGLAVIDFLIQMSAFVALLSLDSKREQASR 806

DB 764 IMLSSVTEVVAFGALGALVMPAVRNFALYAAGSVLFGALMQCTVFSAMTLDLRRSESMR 823
QY 807 LDVCCCVCKQOEUPPPG-----QGEGLLGFQKAYAPFLHWHITRGVLLFLALFGV 859
DB 824 IDCFCPCIRLR--PPIGLYDNEAPSRBEGVMKFMRTVYAFSLRHRHVKQLVAVFGGLFLA 891
QY 860 SLYSMCHISVGLDOELALPKDSVLLDYFLFLANRYPEVGAPVVPVTTTLYGNFSEAGMNAI 919
DB 882 AIIQIHIISLGLDQRLALPSESHLVFYFNDVDSYLDVGPVPVTE--GGDPSRHGQOQL 940
QY 920 CSS-AGCNFFSFTQKIYATPEPEQSYLAIPASSWVDFFIDMLTPS--SCCR-----L 969
DB 941 CGRFTTCLLSVANTLEAERKRPDSFSFIASPPAAWIDDFLOWNTPTFESCCRRRRDPTI 1000
QY 970 YISGPNKDFC-----PSTVNSINCLKNCMSITMGSVRSVPEQFHKYLPWFLNDRP 1020
DB 1001 FCSPKDSERLCRPFEGQKWDSTMDGL-----PEGEDFMKYLQWLSIPT 1045
QY 1021 NIKCPKGLAAYSTVNL--TSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITAD 1079
DB 1046 NDECPLOGQAPYAAVKLVSSNTTVAASHPTVYHPLKSOADFINALAAARRISDDIT-- 1103
QY 1080 LRKVPGTDPAPFVPFYTITNVFYEQYLTILPEGLFMLSCLVPTTFVAVSCLLGLDLRSL 1139
DB 1104 -----HRTGVRVPFYSLVFVFPDQYEHIIISMAIQVLFVAVLVITSTLLG-SWRTGG 1156
QY 1140 LNLISVILVDTVGFMALWDISVNAVSLNLVSAVMSVEFVSHITRSF----- 1189
DB 1157 TVTFTCALAVINVMGVNGYWGISLNAISLVNLISLGIAVEFCSHARAPMGAGSGPLD 1216
QY 1190 AISTKPTWLRRAKATISMGSAVPAVAMTNLPGLVGLGAKAQLIQIFFRMLLTLTL 1249
DB 1217 KLEGRKEDERATWALVDVGFSPVSGITMTKLIGISVLAITRSKLVLEVYFRMWLSLILS 1276
QY 1250 GLLHGLVFLPVILSYVG 1266
DB 1277 GALHGLVLLPVLLSYLG 1293

RESULT 29
Q4WNG5 ASPFU PRELIMINARY; PRT; 1273 AA.
ID Q4WNG5
AC Q4WNG5
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Patched sphingolipid transporter (Ncrl), putative.
GN ORFNames=Afu609980;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H. Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Gobie A.,
RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Jimenez J.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimeenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mounya I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penava M.A., Perteau M., Price C., Pritchard B.H., Quail M.A.,
RA Rabinowitz E., Rawlin N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Rensing C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,

CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL: AABX01000767; BAA26850.1; -: Genomic DNA.

DR GO: GO:0016021; C: integral to membrane; IEA.

DR GO: GO:0016020; C: membrane; IEA.

DR GO: GO:0008158; F: hedgehog receptor activity; IEA.

DR InterPro: IPR000152; Asx hydroxylase.

DR InterPro: IPR003392; Patched.

DR InterPro: IPR00731; SSD 5TM.

DR Pfam: PF02460; Patched; 1

DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.

DR PROSITE: PS01556; SSD; 1.

KW Hypothetical protein.

SQ SEQUENCE 1162 AA; 127777 MW; 6776A78BB9B9619 CRC64;

Query Match 21.7%; Score 1500; DB 2; Length 1162;
 Best Local Similarity 31.4%; Pred. No. 5e-97;
 Matches 383; Conservative 229; Mismatches 437; Indels 170; Gaps 35;

QY 123 LCHHTCSPNQSLFNVTRVAQGLPQAVVAYEAFYQHSFAEQSYDSCSRVRVPAAT 182

DB 1 MECTCTSPNQSLFNVTKTIEKKGL--VTELDQLISEEYGTGFYNSCKDVKEGPTNS 58

QY 183 LAVGTMGVYGALCNAQWMLNFGDTGNGLAPLDTITHLLEPGQAVSGIOPLENGVAR 242

DB 59 RAM----DLIGGAKNYTQLLKFLQERPGSGPFOINF----PVEYABPDMKPLPMKPKK 110

QY 243 CNEOGDDVATSCODCAACPAIARPOALDSTFFYLGOMPGSLVLIILCSFPAVVTILL 302

DB 111 CNDE--DNFRCAVDCPEICPTLPDVEQAGC--HVGALPCLSFASILTYSVILFISIAA 167

QY 303 VGRFV-----APARDKSK-----MVD-PKKGTSLSDKLSFSTH 334

DB 168 VGVHVAWEKHAERSERLLTDAAPDEDEGDTQNVAMIDRQK-----TYIIN 219

QY 335 TLGGQFFQCGWTFVSWPHTLIVLSPVVALAAGLVFTETDTPVELWSAPNSOARSEK 394

DB 220 TWCDSAFSLKGVAAFTPAITVTSILIASILSLGWFFHELEKPNARLWVSPSPAEEK 279

QY 395 AFHDHGFPGFTNQVILTAPNRSYRVDLSLLGPNFSGIILDLLELLELQERLHL 454

DB 280 AFPDHFHGFAYAEKVFELVNDTPQS-----GP---GPVLSRDTLLWMDVKSVAAL 328

QY 455 QVMSPEAQKNISLDQICVAPLNPDNTSLYDCCINSILQYFQNNRTLLLTANQTLWGQTS 514

DB 329 K-----GSNYSGFQDLCKPTGD-----ACVQSVAAVYFQDD-----PD 363

QY 515 QVD---WKDHFLYCANAPLTFKDGFTALALSCMADYGAPVFPFLAIGGY-KGKDYSEBAL 570

DB 364 SVDPTWQSTLRTCAASP-----VECRPAYGQPLDPSMILGGYPEGGNVAEASAM 413

QY 571 TMTFSLNNYPAGDPRLAQAQLWEAFLEEMRAFO--RRWAGMFQVTPFAERSLDEINRT 628

DB 414 TVTWVLIINPSENSEPVDRAMDMEVALKRLLEVEQDEAKERGL-RLSFSTETISLEELNKS 472

QY 629 TAEDLPPIATSYIVIFLYISLALGYSVNSR-----VMVDSKATLGLGGVAVVLGAV 680

DB 473 TWTDAKIVISIIWFLVSLALGASTLTTFKDLINPVALSVSEKFTLGIVGIVILMSI 532

QY 681 MAAMGFFSYGLSRSLVILQVVPFLVSLVGADNIFIVLEYQL-PRRPGEPREYHIGRA 739

DB 533 TASIGLFWAGLRATLIIVDVIPFIVLAVGVNDIFLIVHEPERVNVSPDDMVBARISRA 592

QY 740 LGRVAPSMLLCSLSAICFFGLATFMPAVRTFALTSGTGLAVILDFLQMSAFVALLSIDS 799

DB 593 LGRMGPSILFSALTETASFGALGAFVGMFAVRNFAIYAAGVAFINAILQVTFVSVLTINQ 652

QY 800 KEQEASRLDVCCKVQKQ-----ELPPPGQ-----EGLLGFFQKAVAPFLHWT 845

DB 653 IRVEDSRADCFPCIQIKGARVHLANGAGPAPVYLEAPEESYLQOQFIRKVRAPRLGKKT 712

QY 846 RGWVLLFLALPGVSLYSNCHISVGLDQELALPKDSYLLDYPFLNRYFEVGAAPVYFVT 905

DB 713 KAVIMIEFLGVFAAGVALIPEVKGLDQORVAIPDDSYLIPFNDLYEYLTGPPVYFVTR 772

QY 906 LGYNFSSBAGWNAICSS--AGCNNEFTQIOVATEFPBQSYLAIPASSWDDDFIDWLTPS 964

DB 773 -EFNATDRAQQQKVCARYTTCEQMSLNILSOERKRTVSYISTPTASWIDDFPWLNPE 831

QY 965 S--CCRLYISGNKDKFCPSTVNSLNCUKN---CMSITMGSVRPSVEQPHKYLPMFLNDR 1019

DB 832 NERCC-----MERRRP-----CFANRTPAWNITLSGM-PEGDEFVYVYLKFLSAP 875

QY 1020 PNICPKGGLAAYSTSVNLTSQGVL-ASRPMAYHKPLNSQDYTEALRAARELANITA 1078

DB 876 TNEDCPLGQASYSAGVAVLSDRDTIPASHFRTSHIPLRSQBDFTDAYAAARRIANEISA 935

QY 1079 DLKRVPGTDAPEFPPYPTITNVFEYOYLTILPEGLFMLSCLVPTFVAVSCLLGLDLRS 1138

DB 936 -----ETGLDVFPYSVFFVFPQYASIVSLTGLGSAIGIIFIIVASVLGSLVTA 987

QY 1139 LNLISIVMILVDTVGFNALWDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWL 1198

DB 988 VVS-PTVMAIVDIIGAMAVGVSILNAVSLVNLIIICVIAVEFCAHIAARAFMPSR-SCM 1045

QY 1199 ERAKE-----ATISMGSAVFAGVAMTNLPGIILVGLAKAQLQIPIFFRLNLIT 1247

DB 1046 ERAKNRFRGRDARAWTALSNGVSGVSGITVTKLIGVFLGFTSKRIFEIYVRIWVALV 1105

QY 1248 LLGLLHGLVLPVILSYVG 1266

DB 1106 IPAATHALVLPVALSLVG 1124

RESULT 36

Q59ZV0 CANAL PRELIMINARY; PRT; 1256 AA.

AC Q59ZV0;

DT 10-MAY-2005 (TrEMBLrel. 30, Created)

DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)

DE Potential sterol homeostasis protein.

OS Name=NCR1; ORFNames=CAO19.7242;

GN Candida albicans SC5314.

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitospozic Saccharomycetales; Candida.

ON NCBI_TaxID=237561;

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SC5314;

RX PubMed=15123810; DOI=10.1073/pnas.0401648101;

RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,

RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,

RA Davis R.W., Scherer S.,

RT "The diploid genome sequence of Candida albicans.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334 (2004).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SC5314;

RA Roberts J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,

RA Jones T., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,

RA "Annotation of the Genome of Candida albicans.";

RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AACQ0100092; BAK96004.1; -: Genomic DNA.

SQ SEQUENCE 1256 AA; 141103 MW; 89AF8C4E2ABF5286 CRC64;

Query Match 21.2%; Score 1461.5; DB 2; Length 1256;

Best Local Similarity 28.4%; Pred. No. 3e-94;

Matches 381; Conservative 253; Mismatches 527; Indels 181; Gaps 37;

QY 6 LRGMILLALLRLAQSBEPTTTHOPGYCAFYDECKNPGLSGSLMTLSNVSLSNTPARK 65

Db 13 MRAICLLLLTINLAIAS-VSSSHKPGYCNVTGNGCKKSVFGKPL-----PCAEFVPAVK 65
Qy 66 ITGDHLLILLOKICPRLYTPGNTQACCSAKQLVSLASISITKALLTRCPACSDNFVNLC 125
Db 66 ASQESREKLSKICGKDF----DYICCSPEQIDILESNLKRDVPLISSCPACKKNFYDFFC 121
Qy 126 HNTCSNQSFLINVRVLAQAGQLPAVVAEAFYQHSFASQSDSCSRVRPAAATLAV 185
Db 122 QFSCSPNESQFVEIITKTARDTGKEIVTEINQYVEPGMANQFDFCKNKFSAATN---- 177
Qy 186 GTMCGVYSALCNAORWLNFGDGTGN--GLAPLDITTEHLLPBGQAVSGIQLNEGVARC 243
Db 178 GYMDLIGGGAKNYSQFLKFGDEKPLGGSPYQINFYKLP--ETDGLVLRNEPLRDC 235
Qy 244 NESQDDVATSCQDCAASCPAIRPOALDSTFYLGOMPGSLVLIILILCSFVAVVTILLV 303
Db 236 NDKE----YKCACTDCESCPKLPHAKDOLTKKCTGVLPFCFSFIIL--IWSCMIVLLG 288
Qy 304 GFRVAPAR-----DKSKWVDPKKGTSLSDKLSFSFTHLLGQFFQGW---- 344
Db 289 GYHVYLALEKERRRSIAEDSDDESTMINPLFYAGLGKKGAKQFSSSEIGSKIQDFANI 348
Qy 345 GTWVASPLTILVLSVIPVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHGF 404
Db 349 GYFCSKFPGISIGTSLAVVLLSLGLFKLETDVPKLVSPNDPAYKNOQYFESNFGEW 408
Qy 405 FRTNOVILTAPNRSSRYRYSLLGPKNFGSILDLLELLELQBLRHLQWSPEAQRN 464
Db 409 FRIEQVIVSS-----KDDGPVLNWDIVKWWFKESQLETL-----NEN 446
Qy 465 ISLDQICVAPLNPNTSLYDCINSLLOFYQFNRTLLLTANQTLMGQTSQVDWKDHELY 524
Db 447 VRLSDICEKPL--DET-----CALQSFQYFQGD-----ISGLT-ETNWKSKLQS 488
Qy 525 CANAPLTFKDGATALSCMADYGAVPFPPLAIGYKGYKDYSAEALIMTFSLNYPAGDP 584
Db 489 CVDSP-----VNCUPTFOQLKPNIL---FDSNDISAKAFTVTVLVNSTQEN 535
Qy 585 RLQAQKLWEAFLEEMRAFORRMAGMQVFTFAERSLEDEINRTTAEDLPFATSYIVIP 644
Db 536 YTSNTISYEHFSQKWAADLOTEVPNL-NIAYSTEISLKEELNQSNSTDIKTAISYLVWF 594
Qy 645 LYLISLAGS-----YSSWRVWDSKATIGLGVAVLGVAMAPFSLVGRSSIV 697
Db 595 IYASLALGKLPANLYS-----LVKTRFTLGFSSIIILSVTASVGFPSIIIGURSTLI 649
Qy 698 ILQVPELVLSVGCADNIFIVLEYQRLPR-RGCEPREVHIGRALGRVAPSMLLCSLSEAI 756
Db 650 IAEVIFPLVLAIGDINFLIVHELHVISGNPNLAEVISOALXKHIGPSCFISAVLQVC 709
Qy 757 CFFLALTMPAVRTFALTSGLAVIDLFLQWSAFVALLSLDSKQKQASRLDVC--CCVK 814
Db 710 MFLATSVGNPAVKNFAYTGAGAVLNFSLQMTCTFGLLALQDRLEDNRVDVVPVVTIS 769
Qy 815 P-----QELPPPGQEGLLGFFQKAYAPFLHWTITRGVLLFLALGVSLYSNCHIS 868
Db 770 PIQLQNDDEIDEPVHLEYNFNSRWIGHYAPFLKTKTKPKVITFLVWVGISLSLFPKIQ 829
Qy 869 VGLDQELAPKDSYLLDYFLNRYPEVGAPVFTVTLGYNFSSSEAGMNAICSS-AGCNN 927
Db 830 LGLDQRIAPSKSYLVNYSVVDYLVNVPVFPVVK-DLDYSESNQKICGGPSACDE 888
Qy 928 FSTQKIQVATEPEQSYLAIPASSWDDPIDWLTG--SSCCRLYISG--PNKDKFCPT 983
Db 889 FSLANILEQEFKRSIDISMLSEPASNLDDPFSWLNPDLDQCCRFKSTVFEKTPFCSPN 948
Qy 984 VNSLNCXKMSITMGSVRSVQFHKLP-----WFLND---RPNIKCPKGLAAYSTS 1035
Db 949 APORQC-QSCYL-----NHNPPYDSSMKAFPERDFMFIWNIQEPDQPLGKAHQA 1003
Qy 1036 VNLTSQCVLASRFMAYHKPKNSQDYTEALRAARELANIITADLRKVFQTDPAFVFPY 1095

Db 1004 ISRTTE-KIDSSYFRTSFAPLRGQDEFINAYKSGNNIVKEITKLI-----PSMDVFAY 1055
Qy 1096 TITNVEYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSVMLAVDTVGF 1155
Db 1056 SPFFIFPTQYQNTIVLLTVALLTVAMLIIVYVISTFLJS-SPRAASILTITITAIMINIGV 1114
Qy 1156 MALWDLISYNAVSLINLVSAGVMSVFVSHITRSFA-----ISTKPT 1196
Db 1115 LAWSISLNAVTLVNLVICGVFAVEFTHLTRAYCVPKVMFDPNPAEELYNLNAEPE 1174
Qy 1197 WLER-----AKEATISMGSAVFAVAMTNLPGILVLGAKAQLQIOIFERL 1242
Db 1175 NTRRSSTLSLNAEFRNTKAHNLCSVGGSLISGVTLTKLIGISVLAFTRSQIPEVYFERM 1234
Qy 1243 NLLITLLGLHGLVFLPVILSY 1264
Db 1235 WLSLWISFVHAFVLLPVLISF 1256
RESULT 37
Q551CS D1CDI PRELIMINARY; PRT: 1397 AA.
AC Q551CS;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames:DB0202935;
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Suganag B., Bertram M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal R., Kummerfeld S., Madera M., Konfortov B.A., Rivero P.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pliicher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin D., Lindsay R.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loursched H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL: RAFT01000040; EAL69086.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1397 AA; 156800 MW; 3722FDB55773697F CRC64;

Query Match 21.0%; Score 1452.5; DB 2; Length 1397;
Best Local Similarity 26.6%; Pred. No. 1.5e-93;
Matches 397; Conservative 247; Mismatches 465; Indels 383; Gaps 40;
Qy 23 PYTTHQPGYCAFYDECGKNFELSGLMTLSNVCLSNTPARKITGDHLLILQKICPRLY 82
Db 25 PYNTTLG---CSMYG-----VSSSFVEAKDFPLNNT-----IAPNTQLTH 64
Qy 83 TGPNTQACCSAKQLVSLASISITKALLTRCPACSDNFVNLCNHTCSNQSFLINVTIV 142
Db 65 PEYSIESCCNNTQTLILQTNMLVAGGIFGRCCSCWNLNLCWACSSCPYQKSFVPTKV 124

Qy	143	AQLGAGQIPAVVAYEAFVQHSPAESYSDSCSRVRVPAATAATLAVGTWCGYVGYSALCNAQRW	207
Db	125	-----DNKTDQIISIDIFVLHDPFANGLYNSCRDVH-----	154
Qy	203	LNFGQDTCNGLAPLDITF---HLLPQGAVGSGIOP-----	246
Db	155	-----ANGAQPGFMPTAQLPFTG---VFSAVNPTFKELNWNFNNEGYNDALVPCSEG	205
Qy	247	QGDDVATCSQDCAASC---PAIARPOALDST-----	275
Db	206	-----CSCDQDACLPPSPYKDIGLNDTIQTHLLIEVPHLTIWVMSFLFLLSL	258
Qy	276	-----FYLQMGPSLVLIII---LC-----	292
Db	259	LFGSILWLFKYSSKIRYFGKKFKKPYVMIIILGILFICTVLPIFVTTHQMDGKTASQMW	318
Qy	293	-----SYPAVVTILLVGF-----RVAPARDKSMV-----	317
Db	319	KLDDGKEWNCSPMGVSIYIMITLLLSLIVILACIIVKTSNRNRRILISGGIINOQ	378
Qy	318	-----DPKRGTSLSDKLSFSTHTLLGQFFQCGMGTWASWPLTTLVLVSIPVVALAAGL	370
Db	379	VNDEIQNTNNRRNNNNGSNYSNIQKVFYWGYNITKPIIVLICLLFTTIGIGI	438
Qy	371	VFTELTTDPVLSAPNSQARSEKAFOHQFGPFFRTNQVILTAPNRSYRYSLLGLPK	430
Db	439	ISLEIETDPVKLWSPDSSRAIEKQYFDHDFGPPYRVQEQLIL-IPKQ-----K	485
Qy	431	NFSGILDLDLLELLELQERL-RHLQV-WSPEAQBNISLQICVAPLNPDNTSLYDCCIN	488
Db	486	NLSTIFNYDLFKSLIDIDETQLMKNTSVKLNDSGNRIITLDQLCFEP-----TKRGCLAVE	539
Qy	489	SLLQVQNNRTLLLLLTANQTLMGOTSQVDWKDHFLYCANAPLTKDGTALASCMDYGA	548
Db	540	SITGIFQRDNKL-----NQO---KDNVIGW---FEQCKS-----QLLSPECDMSTGV	581
Qy	549	PVFPPLATGYGKYDYSBAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEMRAFQRMA	608
Db	582	PVNPKIVLGVGWNDS-TLAKAFVTITFLNN---PDSMIDRSMWMEQVWLKIQBISNDPS	637
Qy	609	GMFQVTPTAERSLEDEINRTAEDLPFATSYIVIFLYISALGSYSWSR-----VMVDS	664
Db	638	CPFBSIFNAQRSQDBELAREGNADIPITIVSYFWMFLYISLSGSYYPPFRFNSFFTRS	697
Qy	665	KATLGLGGAVALVGAVMAGMFFSGLIRSSILVLOVPVPLVSLVGADNIIFVLEYORL	724
Db	698	RFALGLSGICIVACSIISVIGCSILRKATLIISEVIPFLVLAIGVDNIIIFLWNTFESL	757
Qy	725	-----PRRCPPEVHIGRALGRVAFPSMLCSLSEIACPFGLGALTMP	767
Db	758	HVASSSSSSISGFNONHHQMPTEE-SFARTLAKVGPSIALASLSSESFAFLGSLTQMP	816
Qy	768	AVRPFALTSLGAVILDFLLQMSAFVALLSLSKQKQESRLDVCCVK-----	814
Db	817	AVKAFSFYASIAVFPDFLLOISAFACCLVLDTKRLQSRVRDCLPCLSLGDNDSLDDPP	876
Qy	815	---PQEL-----PPPGQEG-----	826
Db	877	DNIDQOLNSNNNNNNNNLNNNDNNNNNQBEIENEPLLINGEDDDDDNNANFNVDVAL	936
Qy	827	-----LLLGFFQKAYAPFLHWTIRGVWLLFLA	855
Db	937	NVVDGLNINNESSPLSKRMLKKXKKKTTLAKELFTKYAPFLVHPISKLLIIIFIA	996
Qy	856	LFGVSLYVMCHISVGLDQELALPKDSYLLDYPLFLNRYFVFVGAQPVYFTTILGYNPSBAG	915
Db	997	MLLFSINYAQIPIGLDQKVALPNSYLAQYFTNNMQFLEVGPPEMYIIVKGNYDFSPSI	1056
Qy	916	MNATCSSAGCNFSFTQKIQYATFPEQSILAIAPASSWVDVDFIWLTPSSCCRLYISGN	975
Db	1057	QNQCTVCGCKNNS-----VINTFDNAPFTISPGISSWLDMLQWNTVGCCTAFDNG--	1108
Qy	976	KDFCPSPTVNSLNCNKNCWSITMGVSVPSPVQFHKYLPWFPLNDRPNTKCPKGGGLAAVSTS	1035

Db	1109	--TYCNS--DEPGCDQCFFLEK-SGRPDPSLFKKYLPALFNFTNTDQCPLAGL-AYTGD	1162
Qy	1036	VNLTSDQVQLASPMWYHKPKNSQDYTEALRAARELANITADLRKVPCTGDPAFVFPY	1095
Db	1163	VKFNDSGSIATATRPDAVHTTLRTQNDPINALATSYLLADHSDA-----QVFTY	1210
Qy	1096	TITNVFYEQLTILPEGLFMLSCLVPFAVSCLLGLGLDLSRGLNLLSIYMLVDTVGF	1155
Db	1211	SIFCYFEAVLTIKQVAKLGALLGALGVFPVSLAILMNPVLS-ILVVLGVGMTVDLLAV	1269
Qy	1156	MALWDIISYNAVSLINLVSAGMSVEFVSHITRSPAISTKP-TWLRAKERATISMGSAVPA	1214
Db	1270	MTLWDIISALNAISVNLIMAVGISITFCVHTIASTFIRSPKHYSDKVKFAISEMGANIFS	1329
Qy	1215	GVAMTNLPGLVLGLAKAQLIQIIFPFLNLLITLGLLHGLVLPVLTSVVG	1266
Db	1330	GIPITKLGLGVLAFSKSEFIYYFMYISICLLAAVHGLVLPVLLSIFG	1381
RESULT 38			
Q4PE83_USTWA			
ID	Q4PE83_USTMA	PRELIMINARY;	PRT; 1489 AA.
AC	Q4PE83;		
DT	13-SEP-2005	(TrEMBLrel. 31, Created)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)	
DE	Hypothetical protein.		
GN	ORFNames=U01580.1;		
OS	Ustilago maydis 521.		
OC	Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;		
OC	Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.		
OX	NCBI_TaxID=237631;		
EN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=521;		
RA	Birren B., Nusbbaum C., Abebe A., Abouelleil A., Adekoya E.,		
RA	Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,		
RA	Arachchi H., Armbruster J., Bachanteang P., Baldwin J., Barry A.,		
RA	Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,		
RA	Borowsky M., Boughalter B., Brunache A., Butler J., Calixte N.,		
RA	Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,		
RA	Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,		
RA	David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,		
RA	Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,		
RA	Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,		
RA	Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,		
RA	Girke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,		
RA	Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,		
RA	Hanan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,		
RA	Jaffe D., Jones C., Kamal M., Kamat A., Kamysseles M., Karlsson E.,		
RA	Kells C., Kleu A., Kiser P., Kodira C., Kulbokas E., Labuttl K.,		
RA	Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,		
RA	Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,		
RA	Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,		
RA	Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,		
RA	McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,		
RA	Meisrov J., Mihalev A., Mihova T., Mikelsen T., Mienga V., Moru K.,		
RA	Mozer J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,		
RA	Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,		
RA	Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,		
RA	O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,		
RA	Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,		
RA	Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,		
RA	Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,		
RA	Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sounges C.,		
RA	Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,		
RA	Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,		
RA	Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,		
RA	Towey S., Tsamba T., Tsomo N., Vallee D., Vassiliev H.,		
RA	Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,		
RA	Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,		
RA	Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,		

[3]
 RN NUCLEOTIDE SEQUENCE.
 RA Satou Y.;
 RT "Expressed genes in Ciona intestinalis";
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB210578; BAE06583.1; -; mRNA.
 FT NON_TER
 SQ SEQUENCE 620 AA; 69500 MW; 3032C53A2A0D192 CRC64;
 Query Match 20.1%; Score 1386.5; DB 2; Length 620;
 Best Local Similarity 47.6%; Pred. No. 2.4e-89;
 Matches 293; Conservative 98; Mismatches 187; Indels 37; Gaps 11;
 QY 719 LEYQRLPRRGEPRVHGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTTALTSGL 778
 DB 3 LEYQDERKPOEDLADQIGRVMEVGPVSMILCSLSEALCFPLGALTMPAVRTTALTSGL 62
 QY 779 AVILDFLQMSAFVALLSDSKRQASRLDVCCVKPQELPPPGQGGELGLLGFQKAYAP 838
 DB 63 AIAFDLQITAFVLSLDARTGRNRVDVCCCLIKMEPAEP--NKTYLETFPHKYAP 120
 QY 839 FLLHITRGVLLLFALFGVSLYSIMCHI SVGLDQELALPKDSYLLDYFLFNRYFEVGA 898
 DB 121 VLMDLVRYVMVIGFVGLSCWCTILCTRIITVGLDQDLSPKDSYVLYKYFDYMEKYLDVG 180
 QY 899 PVYFTTLLGYNFSSEAGNAICSSAGCNFFSTQIKIQTATPEPQSYLAIPASSWVDDFI 958
 DB 181 PVYFTKGYNFPADKNASSLICSGAGCDTSLTQISYASQNASYWRITETPAASWYDDIV 240
 QY 959 DMLTP-----SSCRLYISGNKDFKCPSS--TVNSLN-CLKNCMSITWGSVRPSEVQFH 1009
 DB 241 DMLPQGVGGRKSCCRYETHFN--EFCPATTVSKCSFCLKN-----EDYTP--DDFM 290
 QY 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNL-----TSDGQ--VLASRFMAYHKLKNSQDYT 1063
 DB 291 QYLPWFLIDNPGVECNKGHSAYGNVNVNNTSGTGDVVDVDA SYFMAFHSVCIKSVDC 350
 QY 1064 EALRAARELANITADLR-----KVPCTDPAFVFPVTTINNVFVEQVLTILPGLPML 1116
 DB 351 ENLIKARKLADNITLKAANKDGNINLENEEDFEVFPCLYVYVEQVLTAVEOTLQ 410
 QY 1117 SLCLVPTFAVSLGLGLLRSGLNLLSIYMTLVDVTFPMALWDISYNVSLINLVSAG 1176
 DB 411 GCLLIPTFAFSLILGDFYSGIITVITVIMVVDVDTAGLCSLWGVDMNAVSLINLVAIG 470
 QY 1177 MSVEFVSHITRGPASTKPTWLERAKEATISGSAFVAGVAMTNLPGLVLGLAKAQLIQ 1236
 DB 471 LSVEFISHVVRTPSLXTHISKKRVIESMTTGPVAFGVALTNLPGLIIVLNWATAQLIQ 530
 QY 1237 IFFRNLILITLGLLHGLVFLPVILSYVGPVDPNPALEQK--RABEAAVAVMVASCP 1293
 DB 531 IFFRNLVITLGLTGHGLIFLVPVLSYFGPPVKNKALYEEQSELKATQTYMNVKAD 590
 QY 1294 NHPSRVSTADNIYV 1308
 DB 591 QEKQCCKTNKNVYN 605

RESULT 41
 Q618V4_CABR
 ID Q618V4_CABR PRELIMINARY; PRT; 1382 AA.
 AC Q618V4_CABR
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein CBG14473.
 GN Name=CBG14473;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RG The C. briggsae Sequencing Consortium;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CR EMBL; CRACO1000068; CAB68601.1; -; Genomic DNA.
 DR GO; GO:0016021; C:membrane; IEA.
 DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD_5TM.
 DR Pfam; PF02460; Patched; 1.
 DR PROSITE; PS0156; SSD; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1382 AA; 155281 MW; CA269DAB5C9CE527 CRC64;

Query Match 19.7%; Score 1359; DB 2; Length 1382;
 Best Local Similarity 27.8%; Pred. No. 6.3e-87;
 Matches 357; Conservative 253; Mismatches 524; Indels 148; Gaps 36;

QY 74 LQKICPLTYGPNTQACCSAKQLVSLSEASLITKALLTRCPACSDNFVNLHCHTCSPNQ 133
 DB 61 MMEFCPHMEGDN-KUCCTPSQAEGLTKQIAQARHILGRCPSCFDNPAKLWCEFTCSPNQ 119
 QY 134 SLFINVTRVAQLGAGQLPAVVAYEAFYQHS-----FAPQSYDSCSRVRV-- 177
 DB 120 QDFVVISSEMPKEKKE-----GFSEYQPAEAYVNTVEYRLSTEFABGMEFASCKDVTFG 174
 QY 178 -PAAATLAVGTCGVGSGALCNQRLNFOQDGTGNGLA-PLDITFHLLEFQVAGSGIQP 235
 DB 175 QPALRVMCISTPC-----TLTN--WLDFTGTQNLDSLIPINTNFILYDPAKTPQSDRSI 226
 QY 236 -LNEGVARCNESQDDVATCSCODCAAPALRPAQALDSTFYLGOM-----PGSLVLI 289
 DB 227 FMNVNFTGQDSARPGRPACSKSEC--NKEEYAKMINDGKTSQCNVHGIACLAIFV 284
 QY 290 ILCSVFAVVTLLVGFVRPAPADKSKMVDPKKGTSLSKLSFSFTHLLLGQFQGWGTWA 349
 DB 285 MIAFGISLAVLLCVGFVFTSYDEBYTNLRQNAQGESPKNRIRRT-----GAWIH 335
 QY 350 SW-----PLTILVSVIPVVALAAGLVFETLTDPVLSWAPNSQARSEK 394
 DB 336 NFMENNARDIGMAGRNPKSHFFIGCAVLIFCLPGMIYHOESTNVDMWSSPRGRARQEE 395
 QY 395 AFHQHGFPPRTNQVILTAPNRSYRYSVLSLLGPKNPSGILDLLLELEQLERHL 454
 DB 396 TVFNANFORPORYQIIMLLS-NR-EFQNGKLYGP-----VFHKDIPFELFDILNAIKVI 448
 QY 455 QVMSPEAQORNISLODICYAPLNPDNTSLYDCINSLLOYFQ-NRNTLLLLLTANQTLMGQT 513
 DB 449 STKDAEG-RTVTLDDVCYRPMGP-----YDCLIMSPTNYFQARGKENLELKEEIVSD 503
 QY 514 SQVD-----WKDHFYCANAPLTFKDGTLALALSCMADYGAPVFPFLAIGYKGD 563
 DB 504 DAFDYFSSEGTDDMMNHIAACIDQPMSSQK--TKSGLSCFGTYGGPSAPNM-VFGRNTYN 560
 QY 564 YSEAEALIMTFLNNYPAGDPRLAQALWEBAFLEMEAFORRMAGNFQVTF--AERSL 621
 DB 561 YQANSVMTFLVTO--RTEPEIQBAELWEKEFIKCKDYREKSP--KVTFSPFAERSI 615
 QY 622 EDEINRTTAEDLPFATSYIVIFLYISALGSY-----SSSRVWVDSKATILGCVAVV 676
 DB 616 PDEIEKDAKDEIVTVVIALAFILGYVTFSLGRYFVCENQLWS-IIVHSRICLGLSVIIN 674
 QY 677 LGAVMAAGFFSYLGIKRSVLIVQVPLVLSVAGADNIFIFVLEYORL----PRRPGEP 732
 DB 675 LLSFSCSGWIFSMFGIHPVKNALVQFVTVLLGVCRTFMVVKVYQAQLRVSPYMSPDQC 734
 QY 733 EVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTTALTSGLAVILDFLQMSAPV 792
 DB 735 PAIVGVWAGTWPAMFSSSLGCAFPFFIGGFTDLPAIRTFCLYAGLAVILDVVLHCTIFL 794

Db 181 -MC---TSTPCTLTNLEFIGNLNDLNIPIHTKFLYLDPIKTPSPDRSTMTNVTGCD 236
Qy 245 ESQGDVATCSQDCAACPAIARQALDSTFYLGQM-----PGSLVLIILCSVFAVVT 299
Db 237 KSARVGFACSTSEC--NKEEVANLIDLDGKTSQTCNVHGIACLNIFVMLAFIGSLAV 294
Qy 300 ILLVGF-----RVAPARKSKVDPKGTSLSDKLSFSTHTLLGQFF 341
Db 295 LLCVGFVTSYDEDTYTNLRQTSGESPKNRK-----RTGAWI-----HNFMENNA 342
Qy 342 CGWGTWASWPLTILVLSVIPVVALAAGLVFTLTDVPELVWSAPNSQARSEKAFDHOHF 401
Db 343 RDIGMAGRNPKSHFFIGCAVLIFCLPGMIYHKESTNVVDMWSSPRARQAEVFNANF 402
Qy 402 GPFFRTNQVILTPAPNRSSRYRDSLLGPNKPSGILDLDLLELELOBLRHLQWSP 461
Db 403 GRPQRYQIIMLL--SHRDFQSGKLYGP-----VPHKIDIFBELFDILNAIKNISQSDG 455
Qy 462 QBNISLDICVAPLNPDNTSLYDCINSLLQYFQNNRTLLALLTANQT-----L 509
Db 456 -RTITLDDVCRPMGFG---YDCLIMSPNTYFQGNKEHLDKMSKETVSEDDDDAFYF 510
Qy 510 MGQTSQVDWKDFLYCANAPLTFKDGTTALALSCMADYGAFFPFLAIGYKGGQSEAE 569
Db 511 SSEATTDEWMMHMAACIDPHSQK--TKSGLSGMYGPGSPAPNM-VFGKNSHNQAANS 567
Qy 570 LIMTSLNNYPAGDPRLAQAKWEBAFLAEWMAFORMMAGMPQVFTTFAERSLEDRNTT 629
Db 568 IMWTLVQT--RTEPEIQAKELWEKFLKCKEYREKSPKVI-FSFMARSITDIBIEN 624
Qy 630 AEDLPFATSYIVIFLYTSLAGSY-----SSMSRVMDSKATLGLGVAVVGLGVAAM 684
Db 625 KDEIVTWIALAFLIGYVTFSGRYFVCENQLWS--ILVHSRILCGLMSVINLLSSFC 683
Qy 685 GFPSYLGIRSSVLQVFPVLVSGADNIFLVLEY--QRLP-----RRPGPREVHIGRA 739
Db 684 GIFSMFHIHPKVALVQVFFVTLIGVCTFMVKYACQRVSMFMSPDQCEI-VGMV 742
Qy 740 LGRVAPSMLLCSLSBAICFFLGALTPMPARTFALTSGLAVIDLFLQWAFVALLSDS 799
Db 743 MAGTWPAMPSSILGCAFSFFIGFTDLPAIRTFCLYAGLAVLIVLHCTIFLFWMDT 802
Qy 800 KQEQASRLDVCCVKKQELPPQCQEGE---LLG-----FFQKAVAPLLHWI 844
Db 803 QRELNG-----KP-EFFFPYQIKDLGAVLIGRQATDTFTQTFHQVAPFLMRM 853
Qy 845 TRGVULLFLALFGVSLYSWCHISVGLDQELALPKDSYLLDFLFLNRYFEVGAPVYFT 904
Db 854 TRITIGIIFIASFITTVILLSSKISVGFQDSMAFTEKSYISTHRYLDKFDVGPVFFTV 913
Qy 905 TLGYNFSBAGNAICSSAGCNPNFPTQIKOIVATEPPEQSYLAIPASSWDDFDWLT-P 963
Db 914 DGELDHRPDVQNKFTFPFGCSDTSFGNTMNVAVGHTQYLSGEMYNWIDNLEYISRK 973
Qy 964 SSCRLYISGPNKDKCPSTVNSL-----NCLNKM-----SITWGSV---RPSVEQPH 1009
Db 974 SPCKKYVYHDEN--TFCSNTRNKSALDDKACRTCDMDFDVANSYPKSSITWYRPSIEVY 1031
Qy 1010 KYLPWFLNDRPNIKCPKGGIAAYSTSNLTSDQVLASRFMAYHKPL--KNSQDYTEALR 1067
Db 1032 RHLRHLEDTPNSECVFGGRASPDKAISFTSRGIQASQFMTHFKKLSISNSSDFIKAMD 1091
Qy 1068 ARELANIATADLRKVPGTDPAPFVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVS 1127
Db 1092 TARMVSRRLERSI-----DDTAHVFPAYSKIFPFYEQYSTIMPILTQTLFTVWVFGFI 1145
Qy 1128 CLLGLDLRGLNLNLSIVMLVDTVGFEMALWDISYNAVSLNLVSAVMSVEFVSHTR 1187
Db 1146 CVTLGIDVKGACAVICQVSNYTHIVAFMTIFIPVNLASATNLVMSGILIEFSVNVLK 1205
Qy 1188 SFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVILGLAKAQLIQIFPRNLIT 1247

Db 1206 GYACSLRQRAKRAESTVSGSIGPIILSGPVVTVMAGSTMFLSGARLQIITVYFFKFLITI 1265
Qy 1248 LLGLLHCLVLPVTLVSGPDVNPALALEQKRAEEAAVAVMVASCPNHPRSRVSTADNIYV 1307
Db 1266 VSSAHVALIITLPIILAPGSGRHSSTSTNDNDQHDACVLS--PTAESHISNVEEGIL 1323
Qy 1308 N-----HSFEGSIKAGAI 1321
Db 1324 NRPSLLDASHILDPLLKAEGL 1345
RESULT 43
Q12200 YEAST
ID Q12200 YEAST PRELIMINARY; PRT; 1170 AA.
AC Q12200;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, last annotation update)
DE Hypothetical protein YPL006W (Ypl006wp).
GN Name=NCRI; OrderedLocusNames=YPL006W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB972;
RA Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
RA Walsh S.V., Barrrell B.G.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=89057455; PubMed=3143101;
RA O'Hara P.J., Horowitz H., Eichinger G., Young E.T.;
RT "The yeast ADK6 gene encodes homopolymeric amino acid sequences and a
RT potential metal-binding domain";
RL Nucleic Acids Res. 16:10153-10169(1988).
RN (3)
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=91092258; PubMed=2265610;
RA Gerrig S.L., Spencer F., Hieter P.;
RT "The CHL 1 (CTF 1) gene product of Saccharomyces cerevisiae is
RT important for chromosome transmission and normal cell cycle
RT progression in G2/M.";
RL EMBO J 9:4347-4358(1990).
RN (4)
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=97313271; PubMed=9169875;
RA Busey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Brueckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Hunkle-Smith S., Hyman R.W., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle B., Rajandream M.A., Reckmann S.,
RA Rieger M., Riles L., Roberts D., Schaefer M., Scharfe M., Scherens B.,
RA Schramm S., Schroeder M., Sdicu A.-M., Tettelin H., Urrestazu L.A.,
RA Ushinsky S., Vierendeels F., Visser S., Voss H., Walsh S.V.,
RA Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E., Zhong W.-W.,
RA Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";
RL Nature 387:103-105(1997).
RN (5)
RP NUCLEOTIDE SEQUENCE.
RA Jia Y., Cherry J.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN (6)
RP NUCLEOTIDE SEQUENCE.

RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gluchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.,
RT "The genome of the protist parasite Entamoeba histolytica";
RL Nature 433:865-868(2005). The sequence shown here is derived from an
CC -!- CAUTION: EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: AAFB01000798; EAL44130.1; -, Genomic DNA.
DR SEQUENCE 1339 AA; 152565 MW; 3BA116920682B6D7 CRC64;
SQ
Query Match 19.1%; Score 1323; DB 2; Length 1339;
Best Local Similarity 28.2%; Pred. No. 2.1e-84;
Matches 380; Conservative 224; Mismatches 502; Indels 242; Gaps 45;
QY 57 CLSNTPARKITGDHLLQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPAC 116
DB 71 CISEHP-----EFPDNTLGHKT--CDEHSVLDLKEQSTLLDGSFGKCPAC 115
QY 117 SNFVNLHCHNTCSNQSILFINTVRAQLGAGQALPAVAYEAFYQH----SPAQSYDSCS 173
DB 116 YLQNQLMCGFACDPNSARWTEIDDFKHDNTCNEPDGIDVPSLYYLHPLVAQIYDSCK 175
QY 174 RVRVPAATLAVGTWCGVYGSALCNAQRLNFGDTGNGLAPLDTFHLBEGQAVGSGI 233
DB 176 DTEL--SGFAVRSLYPTY-----KEFFQMSIDV-NPVMPLHFYVD--ETGYN-GSKY 222
QY 234 QPLNEGVARCNSQGDVATCCQDCACCAIPA--RQALDSTFVLGMPGSLVLIIL 291
DB 223 TEMH-----KCSQSCHEMCTEFIEDEFTKDYPLDGLTNIITISHL 271
QY 292 CSVFVAVT-----ILLVGR 306
DB 272 FMLFIMVTITVCGGILSYKTFQMKNSPTLLYISVLSGAVTLLGIGLFIINGIIVGFP 331
QY 307 VAPARDKSNVDPKK-----GTSLSKLSFSTHTL----- 336
DB 332 MEGVITMPGWDGNLLGMILLYIGFVGVYGVSVLDYFYVITLTKNKBESDIIDS AKV 391
QY 337 --LGQ-----FF---QWGT---WVASWPLTILVLSVFPVVALAAGLVF 372
DB 392 SILQGESCEEDPNKVDLTNDFLSKMSWYTKFLMKFKW--ILLTVILCCVLCVGVFK 449
QY 373 TELTDPVLSAPNSQARSEKAFHDOHFGPFRTNQVILTAPNRSSYRYSLLGPKNF 432
DB 450 IVFTTDSGLVWPKNQTMKDKNYDEAFGPFRIHQFMLS-----DKLHPG----- 496
QY 433 SGTLDDLLLELLEQLERLHLOV-WSPE---AQRNISLQICYAPLNPNTSLYDCCIN 488
DB 497 NQVLTOPLVQLQOMIDEIRAHVETWDETKETGHTKIDTMDDMCYKRVFGK-----CIIT 551
QY 489 SLIQYFQNNRTLLLTANOTLMQTSQVDWKDHFLYCANAPLTFKDGTLALSCMADYGA 548
DB 552 SVTGYQHDLINKMTENVT-----QYIQNCL-----GNPLATGCADDIGS 592
QY 549 PVFPFLAIGY---KGKDYSEALIMTFSNLNYPAGDPRPRAQAKLWEAFLEMPAFOR 605
DB 593 PVDPHNSLGNVTGDDNNDPMKATILQATFMNQ--PNKTAIEWAELWEEKYLEIL-----D 646
QY 606 RMAGMQVTFPAERSLEDEINRTAEDLPFATSYIVIFLYISLALGS---YSSSRVMV 662
DB 647 RYSLVKVAQAQORSVDDEINRETFTDLVTLVCSYWMFYIYLSLGNSEFYHNYKTVFV 706
QY 663 DSKATLGLGVAVVLGAVMAAMGFSSYLGIRSLVILQVVPFLVLSVGADNIFIP--VLE 720
DB 707 KSRILLGILGIVLFSVFTSAGFFSWINVEATLIITEVIPFLVLAIGVDNIFILNTID 766
QY 721 YQRLPRRPGE--PREVHIGALGRVAPSMCLLSSEACFELGALTPMPAVRFTALTSGL 778
DB 767 EQPMYDXDQGVVPVEKREHLSHVGVGSMMLAISISESLAFFGLYLTSMMPAVQSFSLYAGL 826

QY 779 AVILDFLLQMSAFVALLSLDSKQEQASRLDVCCVKPQELPFPQGE---GLLGFPOKA 835
DB 827 AIFDFILOITVAVNLLCYDKQESHGLDFIPWLQVHDNSLDEQHDFFERGSLIKHLFKY 886
QY 836 YAPFLHMITRGVLLLLFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFFLEAFYFE 895
DB 887 VAVFLSYYPVKIIGLFFIAFFISLNYVPKTMGLGLPQETALPQDSYIQDYFVALT-YLE 945
QY 896 VGAPVYVPTTLGYNFSSEAGMNAICSS--AGCNFNSFTQIQVATEPPEQSYLEIPASSW 953
DB 946 IGPVYIIVRDGATYLNNTQSALCASDNFGCDALSIPNYDAARTIDGTTP-----DW 999
QY 954 VDFIDMLTPSSCCRLYISGNKD-KFCPSPTVNSLNLCKMCSITMGSVRPSVBOFHXYL 1012
DB 1000 IDYFSAQAQKDCCL-----DKDGNICPYDMPNYTECTPCFTNFTDKRVPDPFKYI 1054
QY 1013 PWFANDRN-ICKPKGGLAAYS-----TSVNLTSQ---GOVLASRFMAHYKPKNSQDYTE 1064
DB 1055 NRFLTASINETLCSVNGQAYYVDVVMNHINGTKETDYSYVNVSLRLYHTVLVTQDDFID 1114
QY 1065 ALRAARELANITADLRKVPCTDPAREV--FPYTTITNVFYEQLTILPEGLFMLSCLIVP 1122
DB 1115 AM-----VOAVNISDYMSI-----FDVKTFPPYAHYVYFQYFNIYDLVCMVDCLALAA 1164
QY 1123 TFAVSCILLGLDLRSLNLNLSIVMILVDTVGFMALWDISYNAVSLINLNSAVGMSVEFV 1182
DB 1165 VFIVVMLLM-PDPVIMIVLCVLMCIDLIGIMYLMGVNLNAVSCVNLVMSIGITIEFC 1223
QY 1183 SHITRSPAISTKPTWLERAKEATISGSAVPAGVAMTNLPGILVLGAKAQLIOIFPERL 1242
DB 1224 VHIAHAFSLSSPKTLNDKMDKDAVNMGNVNVGITLTKFLGVIVLSLSSGLIFVITYFRM 1283
QY 1243 NLLITLGLHLGLVPLVILSYGPDVN 1270
DB 1284 YFMMLIFGASHGLFFLPILLSLI-PSVS 1310

RESULT 45

Q750G1 ASHGO PRELIMINARY; PRT; 1178 AA.
AC Q750G1;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE AGL008Wp.
GN Name=AGL008W;
OS Ashbya gossypii (Yeast) (Bremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Bremothecium.
OC NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed:15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Boehmann R., Iuedi P., Choi S., Wang R.A., Flavler A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomycetes cerevisiae genome";
RL Science 304:304-307(2004).
DR EMBL; AGL06820; AAS54482.1; -; Genomic_DNA.
DR AGD; AGL008W; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; P:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1178 AA; 131153 MW; F30AC82F34784F9F CRC64;

Query Match

19.0%; Score 1312.5; DB 2; Length 1178;


```
QY 121 VNLHCHNTCSNQSLFINVTVAQGLGQLPVAVYEAFOHSPAESYSDSCSRVRPAA 180
Db 106 NTIFCHFTCSNQRQFIDIIITDESSDGR-BIVSELNFFLMSWASSFYSCKNKFESAT 164
QY 181 ATLAGTWCYVYGSALCAQWMLNFGQDTGNGL--APLDITFHLLPEQOAVGSGIQPLNE 238
Db 165 N----GYAMDLLGGGASDYSEFMKPLGDKKPKALGGSPPIQINYEY-----ETSDDYISLFE 216
QY 239 GVARCNESQGDVATCSQDCAACAPATARPQALDSTFYLGQMPGSLVLIILCSFAVV 298
Db 217 NVFACS----DPKYKACADCNSCPTLPIK-----KGRCWVAGIPCTSPALL 261
QY 299 TILLVGR-----VAPARKSKMVPKKT-----S 324
Db 262 VLYALAFIAGWGWHIIIFRRRKIKSTNIVEGSEPLDNGNAANSDDIDGLFEVEYATERNA 321
QY 325 LSDKLSFSTHTLLGQFFQMGWGTWASWPLTILVLSVIPV-VALAAGLVFTELTTPVELW 383
Db 322 MNDKISDVLNNTVAACIKN-----PITVLSYSAIFITILLALIVFGDLETDPVNLW 373
QY 384 SAPNSQARSEKAFHDQHGFPFRTNQVIL---TAPNRSSYRYDLSLLGPKNFGSILDLDL 440
Db 374 VNPSPKFKKETHEDANFGPVRTQEVFVNDTGP-----VLSYENLKWFFETEN 423
QY 441 LL--ELLEQERLRLHLOWSPQAQRNISLODICYAPINPDNTSYLDCCINSLLQYFQNNR 498
Db 424 LITENLLSTE-----NISYQDLCFRP-TEDST-----CVIESVTQYFQG-- 461
QY 499 TILLLLTANQTLMGOTSQVDWKDHFYCANAPLTFKDGTLALSCMADYGAFFPFLAIGG 558
Db 462 -----ILPDES---WKQLESCTDSP-----VNCLPSFQQLKTNVL--- 496
QY 559 YRGKDYSEAEALIMFSLNYPAGDPRLAQAQKLWEEAFLEENRAPQORMAGHVFQVTFAE 618
Db 497 FSDENVFESNAFVWLTLLSNH-----TESAKLWEEKL--EQYLLSLGIPDGLRISFNT 548
QY 619 RSLDEINNTAEDLPIFATSYIVIFLYLSLALGYSYNSRWSKATILGGVAVVLG 678
Db 549 MSLEKELN--SSSDIIVSASYLIMFFYASWALKERTGGSRYL-----LGCAGVLIVFS 600
QY 679 AYWAAMGPFYSLGIRSSVILQVWPLVLSVGADNIFIFVLEBYQRL-PRRPGEPREHVIG 737
Db 601 SVIASGSLLSIFGKSTLILAEVIFPLIILAIGDNIPLITHEFDVNSHYSSTIQEKIV 660
QY 738 RALGRVAPSMLLCSLEAICFFLGALTMPAVRTFALTSGLAVIDLDFLQMSAFVALLSL 797
Db 661 LSIKISPSILFSFLCQGGCFLLATWVEMPAVRNFAICAAVALCFNVLQLTYYVCILHF 720
QY 798 DSKRQASRLDVCCCKVQQLPPGQEGLLGFPQKAVAPELLHWITRGVLLFLALF 857
Db 721 YEKYDAANV-----BELTDDNEEGVL-----NRTFLSLDKKRVKVLGVFFSWF 765
QY 858 GVSLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPYFVFTTLGYNFSSEAGMN 917
Db 766 ITSLVLPSPKFLDQTMVAPQNSYLVNYPQDVVEYLVNVPYFVVK-NLDRIKSAQK 824
QY 918 AICSS-AGCNNSFTQKIQVATEFFPQSYLAIPASSWVDDFDWLTP--SSCCRLYISGP 974
Db 825 KICGKTECDYDLSLGNVLELER---LRSSIVEPLANWYDDYMMFLNPELDQCRL---KK 878
QY 975 NKDKFCBSTVNSLNLCKNC-----MSITWGSVRPSVEQFHYLPWFLNDRPNKCPKGG 1029
Db 879 GSEIEIPHPPPRRRC-ETCYKGEWDYDMSGP-PEGDEFMKFYDWIN-APSDPCPLGK 935
QY 1030 AAYSTSVNLTSQGVLASREFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPA 1089
Db 936 SPYSSAI-VYNDNIIITSFRTGHSPLRSQEDYINAVKQARIQVDELN-DL----- 984
QY 1090 FEVFPYITNVFEQYVITLPEGLFMLSCLVPTFAVSCLLGLDLRSLGILLNLSIVML 1149
Db 985 -DVVAYSPFYIFFVQYSTIVSLTAKLITASILLIFFVSWLLGSAITSLSLGI-VTMII 1042
```

```
QY 1150 VDTVGFPMALWDISYNAVSLINLVSAVGMSEVFVSHITRSPAI---STKPTWLRAKEATI 1206
Db 1043 VDLAMMYFPIRLNVLNVLNLLICVLAVEFCIHITRAFTIVPVGVKKDRNSRVIVAMT 1102
QY 1207 SMGSAVFAGVAMTNLPGLVLGLAKAQLIQIPFRNLNLLTLGLLGLVFLPVLISYVG 1266
Db 1103 TIGGSVLKGIWTMTKIIIGISVLALTQSKIFQVFFERMWASLIFVASLHALIFLPVLLSMIG 1162

RESULT 47
Q6FKT2_CANGA
ID Q6FKT2_CANGA PRELIMINARY; PRT; 1193 AA.
AC Q6FKT2;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DB Similar to tr|Q12200 Saccharomyces cerevisiae YPL006w.
GN OrderedLocusNames=CAGL0L0888g;
OC Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OK NCBI_TaxID=5478;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 2001 / CBS 138;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Bolesame A., Boyer J., Fairhead C., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye P., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenn S., Pottier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
DR EMBL; CR380958; CAG62132.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1193 AA; 134734 MW; F9B21267740F476D CRC64;

Query Match 18.4%; Score 1274.5; DB 2; Length 1193;
Best Local Similarity 28.0%; Pred. No. 5.1e-81;
Matches 366; Conservative 240; Mismatches 516; Indels 183; Gaps 51;

QY 10 LLWALLLLRLAQSEPTTTHQPG-YCAFYDECGKNPELSGSLMTLSNSVCLSNTPARKITG 68
Db 12 LIWPFSISIQQPSIQEDKGAQCTIYNGCGK-SIFGSQLPCP-VDDLDFSP-MIDE 68
QY 69 DHLIILLOKICPLRYTGPNTQACCSAKQVLSLEASLSITKALLTRCPACSDNFVNLHCHNT 128
Db 69 ETRELLVSTCGKEWDKDF-ICCTVDQITALKTNLQKAQTISSCPACLENRFLCHFT 127
QY 129 CSPNQSFLINVTVAQLAGQLPAVAVYEAFOHSPAESYSDSCSRVRPAAATLAVGTM 188
Db 128 CSPEQREFVYTKQKSSDGR-SVYDELEVYNNKNTWASSFFDSCKEVKSATN----GYA 182
QY 189 CGVYGSALCAQWMLNFGQDTGNGL--APLDITFHLLPEQOAVGSGIQPLNEGVACRNS 246
Db 183 MDLIGGAKNYTQFLKFLGDKKPKALGGSPFQINF-VYELGDM--DNYRYFNETVYACN-- 237
```



```

OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000078; CAB70180.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; P:hedgehog receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR00731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Hypothetical protein.
FT NON_TER 1 1222
FT NON_TER 1222 1222
SQ SEQUENCE 1222 AA; 137853 MW; DDA709103A13112A CRC64;

Query Match 14.2%; Score 981; DB 2; Length 1222;
Best Local Similarity 25.7%; Pred. No. 3.4e-60;
Matches 332; Conservative 221; Mismatches 575; Indels 164; Gaps 45;

QY 61 TPARKITGDHLLIKICPRLYTGNP-TQACCSAKQLVLEASLSITKALLTRCPACSDN 119
DB 7 TVTRSKNPVYARLEKYCSHLLQEENKAQVCCTQLQLEGMTKRLSVATGILGSCFSDPN 66
QY 120 FYNLCHNTCSNQSIFINVTVAQLGAGQLPAVAYAEFYQHSFABQSYDCSRVRVPA 179
DB 67 FAKLMQCFQSPDQSRFM---KVLSTSGAGKGVKEIEFKVNRDAE-GYTSCKNTW--A 120
QY 180 AATLAVGTMGVYGVGALCNAQWLNFGQDTG---NGLAPLDITHELLEPGAVSGIOP 235
DB 121 SQNLALSFL-----SLGSKVSTENFYGMGRKDPVSNIPMTNFEFTRDKKAMNIPTP 174
QY 236 LNEGVARCMESQDDVATCSQDCQCAACPAIARPOALDSTFYLGMPGSLVLI----- 288
DB 175 -----CHKPANPQTPACGIIDC-----PLESYQLLD-----LSGMEITGVVRVFEHQIKY 218
QY 289 -----IILCSFVAVTILLVGRV-----APARKSKRMVDPKGTSLSDKLSFSTHTLLGQF 340
DB 219 FEWFLRLGCGCTLTMLIPALKYSCHRSPASENCYVDFPGGS-----LEVQFEHLVQY 273
QY 341 FQGMGTWASWPLTILVSVIPVALAAGLVFTELTPDVELWASPNQARSEKAFHQH 400
DB 274 AQ-----VVIHQELKCVGLFVSLCCLGNVWFHSLTHSIDQVSAADGETRRHQKTFIET 329
QY 401 FGPPFRITNQVILT-APNRSYRYDSILLGPKNFSGILDLLELLELLELLELLELLELLELLEL 459
DB 330 FGPTHRIEQVFTFSPSM-----PENFLNQDDTNPFDEMFLINLQNLITVP--- 376
QY 460 EAQRNISLQDICYAPLNPDNTSLYDCCINSLLQYFQNN-RTLLLTANQTLMGQTSQVDW 518
DB 377 QGVSFKTLDDICYRLGKTK-----GCAIMSPTNYEQNNWNTFVNVEDNE-----EDFDY 426
QY 519 KD-----HELYCANAPLFDGTALALSCMADYGAPVFPFLAIGYK-----GKDYSEAE 568
DB 427 NEHNPFTHLKHCFHPFTVK---TPGLSCFGBFGPIDPALVFAGKNNKWHGTEKYTKSR 484
QY 569 ALIMFTSLNNYPAGDPRLAQAALWEAEFLNEARAFQRRMAGMFQVTFTAERSLDEINRT 628
DB 485 TPMIITLLSG-----KNEKAVSWAEFTKMSAYEMKNASP---TFMTESVAHELQAA 536
QY 629 TABDLPIFATSYVIFLYISLALGSY-----SSWSRVWVDSKATLGLGGVAVVLGVAAMAA 684
DB 537 VETDKLVSLACASVLLMINAMIGIYHWPESSLLSAFTHOKLLISTSSVFISVSWCS 596
QY 685 GFPSYLGTRSSVLIVQVPPFLVSVGADNIFVLEVYQBLPRRPEP-----REVH--IGR 738
DB 597 GIYSPGGQHATNAIWLFFVITLIGISRIFTIRTPQSNHNGCHYGHDPDITREMNSTRID 656
```



```
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLTILPGLFMLSCLVPTFAV 1126
Db 179 RSANEIAEQIS-----QNTTAAFPYPSVFYEQYLTIVHDTIFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRLSGLLNLSTVMTLVDVTFGPMALWDISYNVSLINLVSAGVMSVFEVSHIT 1186
Db 231 VFLLGFDDL SAVIVVVTLLLDLDMFGAMYLWNIPLNAVSLVNLVMAVGISVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNL 1244
Db 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFPSKSIKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
Db 351 CVVILGAGHGLVFLPVLSYI 371

RESULT 53
Q66NR4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NR4;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 18;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667365; AAU04384.1; -; Genomic_DNA.
DR EMBL; AY667365; AAU04385.1; -; Genomic_DNA.
DR EMBL; AY667352; AAU04371.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrlflvin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40786 MW; E81FCDD07C34BF83 CRC64;

Query Match 12.6%; Score 873; DB 2; Length 371;
Best Local Similarity 47.0%; Pred. No. 3.2e-53;
Matches 179; Conservative 71; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTLGYNFSSBAGMNAICSSAGCANNFSTOKIYATEPPEQSYLAIPASS 952
Db 1 YLSVGAPYFVVKDQNYTDAAGANQICGGMGNNSLIEQIARMSKMPNYSHIAYPASS 60
QY 953 WDDPFDLWLP--SSCRLYISGPNKDFCPSVNSLNLKNCMSI--TWGSRVPSVEQFH 1009
Db 61 WDDYFDLWKQSSCCRDHNTG--KEDVFCNATVSTSCIA--CRSAQESANQSRPTDPEFM 118
QY 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYREAL 1066
Db 119 KFLPFLNDNPETKCAKGHAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKSKDFIGCL 178
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLTILPGLFMLSCLVPTFAV 1126
Db 179 RSANEIAEQIS-----QNTTAAFPYPSVFYEQYLTIVHDTIFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRLSGLLNLSTVMTLVDVTFGPMALWDISYNVSLINLVSAGVMSVFEVSHIT 1186
Db 231 VFLLGFDDL SAVIVVVTLLLDLDMFGAMYLWNIPLNAVSLVNLVMAVGISVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNL 1244
Db 179 RSANEIAEQIS-----QNTTAAFPYPSVFYEQYLTIVHDTIFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRLSGLLNLSTVMTLVDVTFGPMALWDISYNVSLINLVSAGVMSVFEVSHIT 1186
Db 231 VFLLGFDDL SAVIVVVTLLLDLDMFGAMYLWNIPLNAVSLVNLVMAVGISVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNL 1244
Db 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFPSKSIKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
Db 351 CVVILGAGHGLVFLPVLSYI 371

RESULT 55
Q66NR7_CIOIN PRELIMINARY; PRT; 371 AA.
ID Q66NR7_CIOIN
```

```
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLTILPGLFMLSCLVPTFAV 1126
Db 179 RSANEIAEQIS-----QNTTAAFPYPSVFYEQYLTIVHDTIFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRLSGLLNLSTVMTLVDVTFGPMALWDISYNVSLINLVSAGVMSVFEVSHIT 1186
Db 231 VFLLGFDDL SAVIVVVTLLLDLDMFGAMYLWNIPLNAVSLVNLVMAVGISVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNL 1244
Db 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFPSKSIKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
Db 351 CVVILGAGHGLVFLPVLSYI 371

RESULT 53
Q66NR4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NR4;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 18;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667365; AAU04384.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrlflvin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40786 MW; E81FCDD07C34BF83 CRC64;

Query Match 12.6%; Score 873; DB 2; Length 371;
Best Local Similarity 47.0%; Pred. No. 3.2e-53;
Matches 179; Conservative 71; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTLGYNFSSBAGMNAICSSAGCANNFSTOKIYATEPPEQSYLAIPASS 952
Db 1 YLSVGAPYFVVKDQNYTDAAGANQICGGMGNNSLIEQIARMSKMPNYSHIAYPASS 60
QY 953 WDDPFDLWLP--SSCRLYISGPNKDFCPSVNSLNLKNCMSI--TWGSRVPSVEQFH 1009
Db 61 WDDYFDLWKQSSCCRDHNTG--KEDVFCNATVSTSCIA--CRSAQESANQSRPTDPEFM 118
QY 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYREAL 1066
Db 119 KFLPFLNDNPETKCAKGHAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKSKDFIGCL 178
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLTILPGLFMLSCLVPTFAV 1126
Db 179 RSANEIAEQIS-----QNTTAAFPYPSVFYEQYLTIVHDTIFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRLSGLLNLSTVMTLVDVTFGPMALWDISYNVSLINLVSAGVMSVFEVSHIT 1186
Db 231 VFLLGFDDL SAVIVVVTLLLDLDMFGAMYLWNIPLNAVSLVNLVMAVGISVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNL 1244
Db 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFPSKSIKVFYFRMYL 350
```

```
AC Q6NR7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-FEB-2005 (TrEMBLrel. 28, Last sequence update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 15, West coast 16, Japan 2, Japan 3, Japan 4, and
RC West coast 6;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667362; AAU04381.1; -; Genomic_DNA.
DR EMBL; AY667363; AAU04382.1; -; Genomic_DNA.
DR EMBL; AY667368; AAU04387.1; -; Genomic_DNA.
DR EMBL; AY667369; AAU04388.1; -; Genomic_DNA.
DR EMBL; AY667353; AAU04372.1; -; Genomic_DNA.
DR EMBL; AY667370; AAU04389.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrflvin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40787 MW; D05762F36018549 CRC64;

Query Match 12.6%; Score 869; DB 2; Length 371;
Best Local Similarity 46.7%; Pred. No. 6.1e-53;
Matches 178; Conservative 72; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVVKDQNYTDAAGANQICGGMGNNSLIETQIARMSKPNYSHIAYPASS 60

QY 953 WDDFDIDLTP--SSCRLYISGNPKDKFCPSTVNSLNCNKMSI--TWGSRVPSVEQFH 1009
DB 61 WDDYDFDLKPKQSSCCRDHNTG--EEDVFCNATVVTSCIA--CRSAQESANQSRPTDFEM 118

QY 1010 KYLPWFLNDRPNKICPKGGLAAYSTSVNLTSQG---QVLASRFMAYHKPLKNSQDYTEAL 1066
DB 119 KFLPWFNLNDNPETKCAKGHAAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKTSKDFIGCL 178

QY 1067 RAARELAANITADLRKVPGTDPAPFVFPYTTITNVFYEQLTTLPEGLFMLSCLVPTPAV 1126
DB 179 RSANKIAEBS-----QNTTAEVFPYSVFYFYEQLTIVHDTIFNLGVSLLAIFVV 230

QY 1127 SCLLGLDLRSLNLISVMILVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHIT 1186
DB 231 VFLLGFDLLSAIVVVTLLILLDMFGAMYLWNIPNLNAVSLVNLVMAVGISVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGSS--AVFAGVAMTNLPGILVLGLAKAQLIQIFPRML 1244
DB 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFSKSIQFKVFYFRMYL 350

QY 1245 LITLLGLLHGLVFLPVLISV 1265
DB 351 CVVVLGAGHGLVFLPVLISYI 371

RESULT 56
Q6NS4_CIOIN
ID Q6NS4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q6NS4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-FEB-2005 (TrEMBLrel. 28, Last sequence update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 1;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667362; AAU04381.1; -; Genomic_DNA.
DR EMBL; AY667363; AAU04382.1; -; Genomic_DNA.
DR EMBL; AY667368; AAU04387.1; -; Genomic_DNA.
DR EMBL; AY667369; AAU04388.1; -; Genomic_DNA.
DR EMBL; AY667353; AAU04372.1; -; Genomic_DNA.
DR EMBL; AY667370; AAU04389.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrflvin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40787 MW; D05762F36018549 CRC64;

Query Match 12.6%; Score 869; DB 2; Length 371;
Best Local Similarity 46.7%; Pred. No. 6.1e-53;
Matches 178; Conservative 72; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVVKDQNYTDAAGANQICGGMGNNSLIETQIARMSKPNYSHIAYPASS 60

QY 953 WDDFDIDLTP--SSCRLYISGNPKDKFCPSTVNSLNCNKMSI--TWGSRVPSVEQFH 1009
DB 61 WDDYDFDLKPKQSSCCRDHNTG--EEDVFCNATVVTSCIA--CRSAQESANQSRPTDFEM 118

QY 1010 KYLPWFLNDRPNKICPKGGLAAYSTSVNLTSQG---QVLASRFMAYHKPLKNSQDYTEAL 1066
DB 119 KFLPWFNLNDNPETKCAKGHAAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKTSKDFIGCL 178

QY 1067 RAARELAANITADLRKVPGTDPAPFVFPYTTITNVFYEQLTTLPEGLFMLSCLVPTPAV 1126
DB 179 RSANKIAEBS-----QNTTAEVFPYSVFYFYEQLTIVHDTIFNLGVSLLAIFVV 230

QY 1127 SCLLGLDLRSLNLISVMILVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHIT 1186
DB 231 VFLLGFDLLSAIVVVTLLILLDMFGAMYLWNIPNLNAVSLVNLVMAVGISVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGSS--AVFAGVAMTNLPGILVLGLAKAQLIQIFPRML 1244
DB 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFSKSIQFKVFYFRMYL 350

QY 1245 LITLLGLLHGLVFLPVLISV 1265
DB 351 CVVVLGAGHGLVFLPVLISYI 371

RESULT 56
Q6NS4_CIOIN
ID Q6NS4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q6NS4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-FEB-2005 (TrEMBLrel. 28, Last sequence update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 1;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
```

```
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 8;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667355; AAU04374.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrflvin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40821 MW; EA1DAF58DFB98D95 CRC64;

Query Match 12.6%; Score 869; DB 2; Length 371;
Best Local Similarity 46.7%; Pred. No. 6.1e-53;
Matches 178; Conservative 71; Mismatches 114; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVVKDQNYTDAAGANQICGGMGNNSLIETQIARMSKPNYSHIAYPASS 60

QY 953 WDDFDIDLTP--SSCRLYISGNPKDKFCPSTVNSLNCNKMSI--TWGSRVPSVEQFH 1009
DB 61 WDDYDFDLKPKQSSCCRDHNTG--EEDVFCNATVVTSCIA--CRSAQESANQSRPTDFEM 118

QY 1010 KYLPWFLNDRPNKICPKGGLAAYSTSVNLTSQG---QVLASRFMAYHKPLKNSQDYTEAL 1066
DB 119 KFLPWFNLNDNPETKCAKGHAAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKTSKDFIGCL 178

QY 1067 RAARELAANITADLRKVPGTDPAPFVFPYTTITNVFYEQLTTLPEGLFMLSCLVPTPAV 1126
DB 179 RSANKIAEBS-----QNTTAEVFPYSVFYFYEQLTIVHDTIFNLGVSLLAIFVV 230

QY 1127 SCLLGLDLRSLNLISVMILVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHIT 1186
DB 231 VFLLGFDLLSAIVVVTLLILLDMFGAMYLWNIPNLNAVSLVNLVMAVGISVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGSS--AVFAGVAMTNLPGILVLGLAKAQLIQIFPRML 1244
DB 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFSKSIQFKVFYFRMYL 350

QY 1245 LITLLGLLHGLVFLPVLISV 1265
DB 351 CVVVLGAGHGLVFLPVLISYI 371

RESULT 57
Q6NT1_CIOIN
ID Q6NT1_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q6NT1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 1;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
```

```

RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667348; AAU04367.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrflavin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 371 371
SQ SEQUENCE 371 AA; 40863 MW; B3745263A6AB8B5 CRC64;

Query Match 12.5%; Score 867; DB 2; Length 371;
Best Local Similarity 46.7%; Pred. No. 8.4e-53;
Matches 178; Conservative 71; Mismatches 114; Indels 18; Gaps 7;

QY 893 YFEGAPVYFVTTLGYNFSEAGMAICSSAGCNFSFTQKIQYATFEPPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVKDQGYNTDAAGANQICGGMGNNLSLIEQIARMSKMNYSIAYPASS 60

QY 953 WDDFDLWTP--SSCRLYISGNKDKFCPSTVNSLNCNKMSI--TWGSVRPSVEQPH 1009
DB 61 WLDDYFDWLKPOSSCCRHDNTG--BEDVFCNATVWSTSCIA--CRSAQESANQSRPTDPEM 118

QY 1010 KYLPWFLNDRPNKCPKGLAAYSTSVNLTSDG---OVLASRFMAVHKPLKNSQDYTEAL 1066
DB 119 KFLPFLNDRPNKCPKGLAAYSTSVNLTSDG---OVLASRFMAVHKPLKNSQDYTEAL 178

QY 1067 RAARELANITADLRKVPCTDPAFEVFPYITITNVFYEQYLTILPEGLFMLSCLVPTPAV 1126
DB 179 RSANETAEQIS-----QNTTAEVFPYVSVFVEQYLTIVHTDITFNLGSLAIFV 230

QY 1127 SCLLGLDRSGLLNLSIMILVDTVGFMAWDSYNVSLINLVSAGMSVFEVSHIT 1186
DB 231 VFELGDFPLSAIVVVTILLDMFGAWLWNLPLNAVSLVNLVMAVGISVFECAHIT 290

QY 1187 RSFAISTKTWLERAKEATISNGS--AVPAGVAMTNLPGILVLGLAKAQLIOIPFRLNL 1244
DB 291 RAPALSQRITVARAEALAEISVSVLSGITLTKFVGIVILAFSKSIQPKVYFRMYL 350

QY 1245 LITLGLLGLVPLPVLISYV 1265
DB 351 CVVVLGAGHGLVPLPVLISYI 371

RESULT 58
Q9C8F0_ARATH PRELIMINARY; PRT; 524 AA.
AC Q9C8F0_ARATH PRELIMINARY; PRT; 524 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein T8D8.3 (Fragment).
GN Name=T8D8.3;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025815; AAG51318.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
KW Hypothetical protein.

FT NON_TER 1 1
SQ SEQUENCE 524 AA; 58105 MW; 4414D9C5D4FD0333 CRC64;

Query Match 12.4%; Score 858.5; DB 2; Length 524;
Best Local Similarity 38.1%; Pred. No. 5.3e-52;
Matches 188; Conservative 94; Mismatches 152; Indels 59; Gaps 13;

QY 822 GQGE---GLLGGFFQKAYAPFLHMTTRGVVLLF--LALFGVSLYSMCHISVGLDQELA 876
DB 24 GVGQRKAGLLTRYMKVHAPVLSHWIVKIVVIAFPFGLAWAGIALST--RIEPLGSEQIV 81

QY 877 LPKDSYLDYFLFNRYFEVGAPVFTVTLGVNFSEEA--GMNAICSSAGCNFSFTQKIQ 935
DB 82 LPQDSYLOGYFNNISYLRIGPPLFYVLK--NTNYSSESRTNQLCSINKCNPSLILNEIA 140

QY 936 YATEPEQSYLAIPASSWDDFDLWTPSS--CCRLYISG---PKNCK--PCPSTVNSLN 988
DB 141 RASLTPELSYIAKPAASLDDDFLWLSPEAFCCCRKFTNGTFCPPDDQPPCPGQTS CG 200

QY 989 CLKNCMSIT-----MGSVRPSVEQFHKYLFWFLNDRPNIKCPKGLAAYSTSVNLT-- 1039
DB 201 LSEVKDCTTCPRHADLSSDRFTTQFKEKLPWFLNALPSADCAKGGHAYSSVDLQGY 260

QY 1040 SDQVLAERFMAHYHPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYITN 1099
DB 261 ANGIICASSERTYHTPLNKQVDFVNSMRAAQBSAKVSRSLK-----MBIYPYSVFY 312

QY 1100 VFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDRSGLLNLSIMILVDTVGFMAW 1159
DB 313 MFEQYLDLWKTALNLSIAAFAV--CLITCSFWSAISILLVIAIIIDLLGVMAVF 371

QY 1160 DISYNAVSLINLVSAGMSVFEVSHITRSFAISTKTPTWL----- 1198
DB 372 HQLNALSVNLSVGLAVFCVHTHAPSVPISVCTSIITHKQFAIVLTIVRNAQIST 431

QY 1199 ----BRAKEATISMSGSAVAGVAMTNLPGILVLGLAKAQLIOIPFRLNLTLLGLH 1254
DB 432 GDRNHRMKEALGGMGASVFGITLTKVGVILGFSRSEVFVYVYFKMYLALVLLGLFLHG 491

QY 1255 LVFLPVLISYVCP 1267
DB 492 LVFLPVLISYVCP 504

RESULT 59
Q66NS0_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NS0_CIOIN PRELIMINARY; PRT; 371 AA.
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Clona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cloniidae; Cliona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=West coast 12;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667359; AAU04378.1; -; Genomic_DNA.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
FT NON_TER 1 1
FT NON_TER 371 371
SQ SEQUENCE 371 AA; 40633 MW; 168F3A36B45EEA3 CRC64;

Query Match 12.4%; Score 854; DB 2; Length 371;
Best Local Similarity 46.5%; Pred. No. 7e-52;

```

Matches 177; Conservative 70; Mismatches 116; Indels 18; Gaps 7;
QY 893 YFEVGAIVYFVTLGYNFSSBAGNAICSSAGCNFFSTQKIQVATEPEOSYLAIASS 952
DB 1 YLSVGAPYFVVKDQNTYDAGANQICGCGCNCNNSLIEQIARMSKPNYSHIAYPASS 60
QY 953 WYDDFDMLTP--SSCCRLYISGPNKDKFCPSVNSLNCNKMSI--TWGSRVRSVEQFH 1009
DB 61 WDDYDFDLKQSSCCRDHNTG--EDVECNATVSTSCIA--CRSAQESANQSRTPDDEFM 118
QY 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNL2SDG---QVLASRFMAYHKPLKNSQDYTEAL 1066
DB 119 KFLPWFNDNPTKCAKGCHAAAGTSVKVDEGKSRVGATSFMAHYTLTKTSKDFIGCL 178
QY 1067 RAARELANIATDLRKVGTDPAPRVFPYTTTNNVFEQYLITLPGFLMISLCVPTFAV 1126
DB 179 RSANIEAQIS-----QNTTAEVFPYFVFEQYLTIVHTDITFNLGVSAAICGG 230
QY 1127 SCLLGLDLRSGLNLSIVMLVDTGFMALWDISYNAVSLINLVSAGVMSVRFVSHIT 1186
DB 231 VFLLGFDLLSAIVVVILLILDLDFGAMYLWNIPLNAVSLVNLVMAVGISVECAHIT 290
QY 1187 RSFAISTRTPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIPFRLNL 1244
DB 291 RAFALSQRITRVARAEALAEATIGSSVSLSGITFTKFGVIVILASSKQIFKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
DB 351 CVVVLGAGHGLVFLPVLSYI 371

RESULT 60
QSTQHS ANOGA
ID QSTQHS ANOGA PRELIMINARY; PRT; 601 AA.
AC QSTQHS;
DT 01-FEB-2005 (TreeBLrel. 29, Created)
DT 01-FEB-2005 (TreeBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TreeBLrel. 29, Last annotation update)
DE ENSANGP0000025743 (Fragment).
GN ORFNames=ENSANG00000007926;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC The Anopheles gambiae Sequence Committee;
RG "Anopheles gambiae re-annotation."
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008964; EAL39646.1; -; Genomic_DNA.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
FT NON_TER 1
FT NON_TER 601
SQ SEQUENCE 601 AA; 66036 MW; 0F665802F4F98BFA CRC64;

Query Match 11.2%; Score 776.5; DB 2; Length 601;
Best Local Similarity 42.1%; Pred. No. 4.2e-46;
Matches 166; Conservative 69; Mismatches 128; Indels 31; Gaps 7;
QY 917 NAICSSAGCNFFSTQKIQVATEPEOSYLAIASSWDDFDLWLPSSCCRLYISGPNK 976
DB 4 NLICGGQYCNLDLSLTQLYIASKQPQSTYLARPASSWLDYIDWSAAGPCKQW----NN 59

QY 977 DKFCPSTVNSLNCNKMSITMGSRVRSVEQFHKYLFWFLNDRPNIKCPKGLAAYSTSV 1036
DB 60 GSFCPHQKSACGA--CNISMTAQKRVPSSFRQYVSFFLEDNPDEACAKAGHAAYGSGV 116
QY 1037 NLTSD-----GOVLASRFMAYHKPLKNSQDYTEALRAARELANIATDLR---KVPQTD 1087
DB 117 KYRPDALAPLYNDVGASYFMAYHTILKSSSDYTEALRSARKISANITSTIHANLRLOGRS 176
QY 1088 PA----FEVFPYTTTNNVFEQYLITLPGFLMISLCVPTFAVSCLLGLDLRSGLNLL 1143
DB 177 EADIQQLVFPYFVFEQYLTWMDPTLKSNGISVLAIFFVFLMGFDIHSLSVVI 236
QY 1144 SIVMLVDTGFMALWDISYNAVSLINLVSAGVMSVRFVSHITRSFAISTRTPTWLERAKE 1203
DB 237 TITWIVINIGGLVHWSISLNAVSLVNLVMAVGISVEFCHLVHSPFAMSVEETREKRAAD 296
QY 1204 ATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIPFRLNLITLLGLHLGLVFLPVLS 1263
DB 297 ALTKWGSVFSGITLTKFGGILVLGFAHSQITQVFFYFRMYLGIVLGAAHGLVFLPVLS 356
QY 1264 YVGPDVNPALALEOKRAEEAAVAVMVASCPNHPS 1297
DB 357 YIG-----VMHRRRRSNDND-----SNVPHPS 379
RESULT 61
PTCL_HUMAN
ID PTCL_HUMAN STANDARD; PRT; 1447 AA.
AC Q13635; Q13463; Q5VZC0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Patched protein homolog 1 (PTC1) (PTC).
GN Name=PTCH; Synonyms=PTCH1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANTS BCC PRO-175 AND
RP PRO-ASN-ILE-815 INS.
RC TISSUE=Lung;
RX MEDLINE=96247324; PubMed=8658145;
RA Johnson R.L., Rothman A.L., Xie J., Goodrich L.V., Bare J.W.,
RA Bonifas J.M., Quinn A.G., Myers R.M., Cox D.R., Epstein E.H. Jr.,
RA Scott M.P.;
RT "Human homolog of patched, a candidate gene for the basal cell nevus
RT syndrome.";
RL Science 272:1668-1671 (1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND VARIANT PRO-1315.
RX PubMed=15164053; DOI=10.1038/nature02465;
RA Humphray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.E.,
RA Howe K.L., Andrews T.D., Searle S., Hunt R.E., Scott C.E., Jones M.C.,
RA Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,
RA Babbage A.K., Babbage S., Bagguley C.L., Bailey J., Banerjee R.,
RA Barker D.J., Barlow K.F., Bates K., Beasley H., Beasley O., Bird C.P.,
RA Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burrill W.,
RA Burton J., Carter C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,
RA Clark S.Y., Clee C.M., Clegg S., Collier R.B., Corby N., Crosier M.,
RA Cummings A.T., Davies J., Dhami P., Dunn M., Dutta I., Dyer L.W.,
RA Earthrowl M.E., Faulkner L., Fleming C.J., Frankish A.,
RA Frankland J.A., French L., Fricker D.G., Garner P., Garnett J.,
RA Ghori J., Gilbert J.G.R., Gibson C., Grafham D.V., Gribble S.,
RA Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,
RA Hammond S., Harley J.L., Harrison E.S.I., Hart E.A., Heath P.D.,
RA Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,
RA Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kerhaw J.K.,
RA Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C., Lloyd D.M.,
RA Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,

RA McLay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,
RA Nordstiek G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,
RA Pelan S., Phillimore B., Povey S., Ramsey Y., Rand V., Scharfe M.,
RA Sehra H.K., Showkeen R., Sims S.K., Skuce C.D., Smith M.,
RA Steward C.A., Swarbrick D., Sycamore N., Tester J., Thorpe A.,
RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,
RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,
RA Young L., Ashurst J.L., Coulson A., Blocker H., Durbin R.,
RA Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,
RA Rogers J., Dunham I.,
RT "DNA sequence and analysis of human chromosome 9";
RL Nature 429:369-374(2004).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA] OF 152-1447.
RC TISSUE=Fetal brain;
RX MEDLINE=96218118; PubMed=8647801; DOI=10.1074/jbc.271.21.12125;
RA Hahn H., Christiansen J., Wicking C., Zaphiropoulos P.G.,
RA Chidambaram A., Gerrard B., Vorechovsky I., Bale A.E., Toftgard R.,
RA Dean M., Wainwright B.J.;
RT "A mammalian patched homolog is expressed in target tissues of sonic
RT hedgehog and maps to a region associated with developmental
RT abnormalities";
RL J. Biol. Chem. 271:12125-12128(1996).
RN [4]
RP VARIANTS BCNS ARG-509; VAL-509; GLN-816 DEL AND TYR-1132.
RX MEDLINE=96438609; PubMed=8940969;
RA Chidambaram A., Goldstein A.M., Gailani M.R., Gerrard B., Bale S.J.,
RA DiGiovanna J.J., Bale A.E., Dean M.;
RT "Mutations in the human homologue of the Drosophila patched gene in
RT Caucasians and African-American nevoid basal cell carcinoma syndrome
RT patients";
RL Cancer Res. 56:4599-4601(1996).
RN [5]
RP VARIANTS BCNS TYR-513 AND ARG-1069.
RX MEDLINE=97136566; PubMed=8981943;
RA Wicking C., Shanley S., Smyth I., Gillies S., Negus K., Graham S.,
RA Suthers G., Haitees N., Edwards M., Wainwright B.J.,
RA Chenevix-Trench G.;
RT "Most germ-line mutations in the nevoid basal cell carcinoma syndrome
RT lead to a premature termination of the PATCHED protein, and no
RT genotype-phenotype correlations are evident";
RL Am. J. Hum. Genet. 60:21-26(1997).
RN [6]
RP VARIANT NBCCS ASP-1438.
RX MEDLINE=98001068; PubMed=9341860; DOI=10.1007/s004390050541;
RA Lench N.J., Telford E.A.R., High A.S., Markham A.F., Wicking C.,
RA Wainwright B.J.;
RT "Characterisation of human patched germ line mutations in naevoid
RT basal cell carcinoma syndrome";
RL Hum. Genet. 100:497-502(1997).
RN [7]
RP VARIANT PRO-1315.
RX MEDLINE=99211211; PubMed=10200051;
RX DOI=10.1002/(SICI)1098-1004(1998)11:6<480::AID-HUMU9>3.3.CO;2-W;
RA Hasenpusch-Theil K., Bataille V., Laehdetie J., Obermayr F.,
RA Sampson J.R., Fritschau A.-M.;
RT "Gorlin syndrome: identification of 4 novel germ-line mutations of the
RT human patched (PTCH) gene";
RL Hum. Mutat. 11:480-480(1998).
RN [8]
RP VARIANTS BCNS SER-376 AND VAL-1083 INS, AND VARIANT BCC TRP-1114.
RX MEDLINE=98281604; PubMed=9620294;
RX DOI=10.1046/j.1523-1747.1998.00222.x;
RA Assterbaum M., Rothman A.L., Johnson R.L., Fisher M., Xie J.,
RA Bonifas J.M., Zhang X., Scott M.P., Epstein E.H. Jr.;
RT "Identification of mutations in the human PATCHED gene in sporadic
RT basal cell carcinomas and in patients with the basal cell nevus
RT syndrome";
RL J. Invest. Dermatol. 110:885-888(1998).
RN [9]
RP VARIANT PRO-1315.
RX MEDLINE=20334946; PubMed=10874314;
RX DOI=10.1002/1098-1004(200007)16:1<89::AID-HUMU18>3.3.CO;2-Z;
RA Dong J., Gailani M.R., Pomeroy S.L., Reardon D., Bale A.E.;
RT "Identification of PATCHED mutations in medulloblastomas by direct
RT sequencing";
RL Hum. Mutat. 16:89-90(2000).
RN [10]
RP VARIANT BCNS PRO-1132.
RX MEDLINE=21154106; PubMed=11231326;
RX DOI=10.1046/j.1523-1747.2001.01279-2.x;
RA Reifemberger J., Arnold N., Kiechle M., Reifemberger G., Hauschild A.;
RT "Coincident PTCH and BRCA1 germline mutations in a patient with nevoid
RT basal cell carcinoma syndrome and familial breast cancer";
RL J. Invest. Dermatol. 116:472-474(2001).
RN [11]
RP VARIANTS SQUAMOUS CELL CARCINOMA MET-829 AND LYS-1242.
RX MEDLINE=21184537; PubMed=11286632;
RX DOI=10.1046/j.1523-1747.2001.01301.x;
RA Ping X.L., Ratner D., Zhang H., Wu X.L., Zhang M.J., Chen F.F.,
RA Silvers D.N., Peacock M., Tsou H.C.;
RT "PTCH mutations in squamous cell carcinoma of the skin";
RL J. Invest. Dermatol. 116:614-616(2001).
RN [12]
RP VARIANTS BCNS PRO-230 AND 505-LEU-ARG-506.
RX PubMed=15459969; DOI=10.1002/humu.9289;
RA Savino M., d'Apolito M., Formica V., Baorda F., Mari F., Renieri A.,
RA Carabba E., Tarantino E., Andreucci E., Belli S., Lo Muzio L.,
RA Dallapiccola B., Zelante L., Savoia A.;
RT "Spectrum of PTCH mutations in Italian nevoid basal cell-carcinoma
RT syndrome patients: identification of thirteen novel alleles";
RL Hum. Mutat. 24:441-441(2004).
CC -!- FUNCTION: Acts as a receptor for sonic hedgehog (SHH), indian
CC hedgehog (IHH) and desert hedgehog (DHH). Associates with the
CC smoothened protein (SMO) to transduce the hedgehog's proteins
CC signal. Seems to have a tumor suppressor function, as inactivation
CC of this protein is probably a necessary, if not sufficient step
CC for tumorigenesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: In the adult, expressed in brain, lung, liver,
CC heart, placenta, skeletal muscle, pancreas and kidney. Expressed
CC in tumor cells but not in normal skin.
CC -!- DEVELOPMENTAL STAGE: In the embryo, found in all major target
CC tissues of sonic hedgehog, such as the ventral neural tube,
CC somites, and tissues surrounding the zone of polarizing activity
CC of the limb bud.
CC -!- PTM: Glycosylation is necessary for SHH binding (By similarity).
CC -!- DISEASE: Defects in PTCH are probably the cause of basal cell
CC nevus syndrome (BCNS) [MIM:109400]; also known as Gorlin syndrome
CC or Gorlin-Goltz syndrome. BCNS is an autosomal dominant disease
CC characterized by nevoid basal cell carcinomas (NBCCS) and
CC developmental abnormalities such as rib and craniofacial
CC alterations, polydactyly, syndactyly, and spina bifida. In
CC addition, the patients suffer from a multitude of tumors like
CC basal cell carcinomas (BCC), fibromas of the ovaries and heart,
CC cysts of the skin, jaws and mesentery, as well as medulloblastomas
CC and meningiomas. PTCH is also mutated in squamous cell carcinoma
CC (SCC). Could also be associated with large body size observed in
CC BCNS patients.
CC -!- DISEASE: Defects in PTCH are a cause of sporadic basal cell
CC carcinoma (BCC) [MIM:605462].
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC -!- DATABASE: NAME=PTCH mutation database;
CC WWW="http://www.cybergene.se/PTCH/ptchbase.html".
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/PTCH100.html".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U59464; AAC50550.1; -; mRNA.
CC EMBL; AL161729; CAH73817.1; -; Genomic_DNA.
DR

DR	GO; GO:0007224; P:smoothed signaling pathway; IGI.	
DR	InterPro; IPR003392; Patched.	
DR	InterPro; IPR004766; Patchedtm_recept.	
DR	InterPro; IPR000731; SSD 5TM.	
DR	Pfam; PF02460; Patched; 1.	
DR	TIGRFAMs; TIGR00918; 2A060602; 1.	
DR	PROSITE; PS0156; SSD; 1.	
KW	Glycoprotein; Receptor; Transmembrane.	
FT	TOPO DOM 1 86	
FT	TRANSMEM 87 107	
FT	TOPO DOM 108 422	
FT	TRANSMEM 423 443	
FT	TOPO DOM 444 458	
FT	TRANSMEM 459 479	
FT	TOPO DOM 480 487	
FT	TRANSMEM 488 508	
FT	TOPO DOM 509 533	
FT	TRANSMEM 534 554	
FT	TOPO DOM 555 563	
FT	TRANSMEM 564 584	
FT	TOPO DOM 585 734	
FT	TRANSMEM 735 755	
FT	TOPO DOM 756 1013	
FT	TRANSMEM 1014 1034	
FT	TOPO DOM 1035 1039	
FT	TRANSMEM 1040 1060	
FT	TOPO DOM 1061 1069	
FT	TRANSMEM 1070 1090	
FT	TOPO DOM 1091 1107	
FT	TRANSMEM 1108 1128	
FT	TOPO DOM 1129 1140	
FT	TRANSMEM 1141 1161	
FT	TOPO DOM 1162 1434	
FT	DOMAIN 424 584	
FT	CARBOHYD 127 127	
FT	CARBOHYD 298 298	
FT	CARBOHYD 335 335	
FT	CARBOHYD 400 400	
FT	CARBOHYD 861 861	
FT	CARBOHYD 986 986	
SQ	SEQUENCE 1434 AA; 159273 MW; A5E9189B633173D0 CRC64;	
Query Match		
Best Local Similarity 22.9%; Pred. No. 2.2e-36;		
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;		
QY	356 LVLSVIPVVALAAGLVFETLTDVPELWSAPNSQARSEKAFHQHGFPRFN-QVILTA 414	
DB	88 LVVGLLIFGAFVGLKAANLETNVEELWVEGVRSRELNTRYQKIGBEAMFNPQLMIQT 147	
QY	415 PNRSSRYDSLLGPKNFGSLDLDLLELE--LQERLRHLQVMSPEAQRNLSLODICY 472	
DB	148 PKEEG-----ANVLTEALLQHLDSALQASRVHYMYN-----ROWKLEHLCY 190	
QY	473 AP-----LNPDNSTLYDCCINSLLQVFNQNRLLLLLTANQTLMGQ----- 512	
DB	191 KSGELITETGYMDQIIEYLYPCLIIITPLDCFEGAKLQSGTA--YLLGKPLPLRWTFDPL 248	
QY	513 -----TSQVDWKDHFLY-----CAN-----APL----- 530	
DB	249 EFLEELKKINYQVDSWEMLNKAEGVGYMDRPLNPDADPCPATAPNKNSTKPLDVALV 308	
QY	531 -----TFKDG-----ALALSCMADYGAPVFPFLAIGYKKG 562	
DB	309 LNGCGGLSKRYMHMOSELIVGTGVKNATKLVSAHALQTMFQMTPKQWYEHFRGY--- 365	
QY	563 DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFQRMAQMFQVTFAPERSLE 622	
DB	366 DY-----VSHINWNE-----DRAAAILEAQRTYVEVHQSVAPNSQKVLPTT-TTLD 414	
QY	623 DEINRTTAEDLPATSYIIVFLYISALGSYSWSRWVDSKATYGLGVAIVLVAVMA 682	
DB	415 DILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKSQAGVLAGVLLVALSVAA 469	

QY	683 AMGFPSVLGIRSSLVILQVVPFLVLSVGDADNIFIFVLEYQRLPRRPPCEPREVHIGRALGR 742	
DB	470 GLGLCSLIGISFNAATQVLPFLALGVDDVFLLAHAFSETQNKRIPIPEDRTGECLKR 529	
QY	743 VAPSMILCSLSEACIFLPGALTEMPAVRTFALTSGLAVLDFLLQMSAFVALLSLDSKQ 802	
DB	530 TGASVALTSISNVTAFMAALIPALRAFSLQAQVVVVVFAMVLLIFPAILSMDLVRR 589	
QY	803 EASRLDVCCC-----VKPOEL-----PPGQGEGL----- 827	
DB	590 EDRLDIFCCFTSPCVSRVIOVEPOAYTEPHSNTRYSPPPPTSHSHAETHITMQSTVQ 649	
QY	828 ----- 827	
DB	650 LRTEYDPTHVYTTAEPARSEISVQPVTVTQDNLSQSPSTSTRDLLSQFSDSSLHCL 709	
QY	828 -----LLGPFQKAYAPFLHLWITRGVLLVFLALFGVSLYSMCHISVGLDELALPK 879	
DB	710 EPPCTKWTLSSFAEKHYAPFLPKAKVWVILLFLGLLVSLVGTTRVRDGLDLTDIVPR 769	
QY	880 DSYLLDYFLFNRYFEVGAPVYFVTTLGYNFSSEAGNAICSSAGCNFSFQKIQVAT- 938	
DB	770 ETRYDFIAAQKYFSF-----YNM-----YIVTKADFPNI 801	
QY	939 -----EFPEQSILAIPASS-----WVDDDFIDWL-----TPSS--- 965	
DB	802 QHLLYDLHLKSFNSVYVNLSEKQLPQMWLHYFEDWLQGLQDAPDSWETGRIMPNYKN 861	
QY	966 -----CCRLYISGPNKDFCPSTVNSLNCMKMSITWGSVPSVPSVEQPHKYL-PWPLN 1017	
DB	862 GSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWVSN 915	
QY	1018 D-----RPNIKCPKGLAAY--STSVNLTSQGVLASRFMAYHKPLKNSQD 1061	
DB	916 DPVAYASQANIRPHRPEWHDK---ADYMPETRLRIPAAEPIEYAQPFYVNLGLRDTSD 972	
QY	1062 YTEALRAARELAANITA-DLRKVFETDPAPEVFFYTTNVFYEQVLTILPGLFMLSICL 1120	
DB	973 FVEALEKRVVICNNVTSGLSSYPNG-----YFP-----LFWEQVISLRHMLLSISVVL 1022	
QY	1121 VPTFAVSCLLGLDLRSLGNLLSVMLLVDTVGFMAWDISYNALINLVSAVMSVSE 1180	
DB	1023 ACTFLVCAVFLNFWPTAGII-VVVALMTVELFGMGLIGIKLSAVPVVILLASVIGIVE 1081	
QY	1181 FVSHITRSF--AISTKPTWLERAKEATISMGSAFAGVAMTNLPGLVGLAKAQLIQIF 1238	
DB	1082 FTVHVALAFLTAIGDKN---HRAMLALEHMFAPVLDG-AVSTLLGLVLMAGSERPDFIVRY 1137	
QY	1239 FFRMLMLTLGLLHGLVFLPVLSYVG--PDVNPALAEQ---KRAEEVAANVAVASCP 1293	
DB	1138 FFAVLAITLVGLNGLVLLPVLSFFGCPCEVSPANGNLRLPTSPSPPPSVVRFAVPP 1197	
QY	1294 NHPRSVS-TADNIYVNHSEFSI-----KGAG 1319	
DB	1198 GHTNNGSDSDSEYSSQTTVSGISEELRQYEAOQAG 1234	
RESULT 64		
QSRJ13 BRARE		
ID	QSRJ13_BRARE PRELIMINARY; PRT; 827 AA.	
AC	QSRJ13;	
DT	01-FEB-2005 (Tremblrel. 29, Created)	
DT	01-FEB-2005 (Tremblrel. 29, Last sequence update)	
DT	01-FEB-2005 (Tremblrel. 29, Last annotation update)	
DE	Novel protein (Fragment).	
GN	Name=OTDARP0000008780; ORFNames=DKEY-61p9.10-001;	
OS	Brachydanio rerio (Zebrafish) (Danio rerio).	
OC	Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Cypriniformes;	
OC	Cyprinidae; Danio.	
OX	NCBI_TaxID=7955;	
RN	[1]	


```

RP NUCLEOTIDE SEQUENCE.
RA Pelan S.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX005238; CAI21193.1; -; Genomic DNA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0008158; F: hedgehog receptor activity; IEA.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS0156; SSD; 1.
FT NON_TER 827 827
SQ SEQUENCE 827 AA; 92019 MW; A7F10494A2C9C041 CRC64;

Query Match 9.2%; Score 636.5; DB 2; Length 827;
Best Local Similarity 24.1%; Pred. No. 5.4e-36;
Matches 233; Conservative 175; Mismatches 371; Indels 187; Gaps 33;

QY 337 LQPFQGGTGWASWPLTILVLSVIPVALAAGLVF-TELTTDPVELMSAP-NSQARSEK 394
DB 12 LSLVPEKLGRLGRHPVVFILLSLVYAAGLGAGFIFLKEREANDIEDQFTPVNGPAKDR 71
QY 395 AFHQHGFGRFTNQVILTPAPRSTRSYRYSLLGPNFSGILDLDLLELLELQRLRL 454
DB 72 EIVAEHFPSPDSFQSLRL-----MSEGTVASLIITDLQGENILTAFAFEILALDRQVKTL 127
QY 455 QVWSPQAQRNLSLDICVAPLNPNTSLVDCCINSLLQYFQNNR-----TLLLTANQT 508
DB 128 Q-----HLGNTFEKLC-AKIIG-----NCVSNVAVLDIIRYNAADISVTIYPINOKT 174
QY 509 LMGQTSQVMDKHDFLYCANAPLTFKDGTAALSCMADYGAPVFPLAIGYKGYDS--- 565
DB 175 FLGTT-----IGGVETQPNSSML 192
QY 566 -BAAALIMTFSLNTPAGDPRLAQAQL-WEEAFLEEMRAF--ORRMAGHGFQVTFABRSL 621
DB 193 KSAKAIRLYYFL-----DEKSKGNADWLEGIQFNSNYTDQEKVS-----VSYFTSVSR 242
QY 622 EDEINRTAEDLPFATSVIVIFLISIALGYSYSSWSRVMWDSKATLGIGVAVVLGAVM 681
DB 243 QNEFEGNTDSVLPFSITTA-----LAINIAV--LSCLRLDCVTKVWVALLGVVVSAGMAVL 297
QY 682 AAMGFPSYLGIRSLVILQVFPFLVSLVGADNIFIVLEYQLRPRRGPPEVHVHIGRALG 741
DB 298 ASFGLLFCGMPFAMTV-GSAFLILGVGVDDFMFIMSSQKTAVDKGV--EFLRAEAYK 354
QY 742 RVAPNMLCLESSEAICFFIGALTTPMPAVRTFALTSLGLAVILDFLQMSAFVALLSLDSKR 801
DB 355 EAGVSITITTLTDVLAFAFYIGLLTPFRSVQSCFWYTSALLFCYLFNITPFGACALNGRR 414
QY 802 QEASRLDVCCVKPQELPPGQEGILLG-----FFQKAYAPELLHWI 844
DB 415 EKSNRHFLTCMTVPK---PSGDAVSCCAGAGFADENTKEHMPMEVFFPKYGGPFLAKVW 471
QY 845 TRGVVLLLFALFGVSLYSMSCHISVGLDQELALPKDSYLLDFLFLNRYFEVGPVYFT 904
DB 472 VKVIVCLIVAGVLAISYICGFQWEGGLDKHLATDGSVADYDREDFSFAGFNWMLV 531
QY 905 TLGVNPF---SSBAGNNA-ICSSAGACNFFSTQIKIYATEFPFQSYLAIPASSWVDPDFIDW 960
DB 532 IKDEHFQVWSPARKSLDLC-----LKNFRDLTMVD-----SEIPLTSLWLDAYM-- 575
QY 961 LTPSFCCRLYISGPNKDFCPSTVNSLNCNMTITWGSVPSVEQPHKYLFWFLNDRP 1020
DB 576 -----XFGQSSFD-----LNNEMI-----PKTQLPAFIN--- 599
QY 1021 NIKCPKGLAAYSTSVNLTSQVLAISRFMAYHKLKNSQDYTEALRAARELANITADL 1080
DB 600 -----RSESHDVHFT--DNNINATRMPIQTVNIKTAIDKMDLNAFRE-AANTCGL 649
QY 1081 RKVPGTDPAPFEVFPYITNVFYEQVLTITLPEGLFMLSLCLVTPFAVSCLLGLDLRSGLL 1140

Db 650 -----ETPVDLIVYHPAPIYFDQYAVIVSNNTQNLVAATCVMLVISLLIHPPLCS-LW 702
QY 1141 NLLSVIMILVDTVGPMALWDISYNAVSLINLSVAVGMSVEFVSHITRSFAISTKPTWLSR 1200
DB 703 VTFSTASIVVGLAGFPMALWDISLDSVSMNLVLCIGFSDVFSAHISYAFVSSEKSSANEK 762
QY 1201 AKBATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFFRNLNLTITLGLHLGVFLPV 1260
DB 763 ATDALHKLGYPIIQG-AVSTIAGVVVLAARSAKSYIPRT-FFKIMFLVILFGALHGVFLPV 820
QY 1261 ILSYVG 1266
DB 821 FLSPFG 826

RESULT 65
PTCL CHICK
ID PTCL_CHICK STANDARD; PRT; 1442 AA.
AC Q90693;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Patched protein homolog 1 (PTCL) (PTC).
GN Name=PTCH; Synonyms=PTC;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus
OC NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96205046; PubMed=8620849;
RA Marigo V., Scott M.P., Johnson R.L., Goodrich L.V., Tabin C.J.;
RT "Conservation in hedgehog signaling: induction of a chicken patched
homolog by Sonic hedgehog in the developing limb.";
RL Nature 384:176-179 (1996).
CC -!- FUNCTION: Acts as a receptor for sonic hedgehog (SHH), indian
hedgehog (IHH) and desert hedgehog (DHH). Associates with the
smoothed protein (SMO) to transduce the hedgehog's proteins
signal.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expression is seen in the embryonic neural
tube, sclerotome, visceral mesoderm, and limb bud.
CC -!- DEVELOPMENTAL STAGE: In stage 10 embryo, expression is seen in
neural tube, and at lower levels in the notochord, epithelial
somites, endoderm and splanchnic mesoderm. At stage 18, PTC is
broadly expressed in the neural tube but excluded from the cells
of the floor plate. At stage 32, expression occurs in the
mesodermal cells of the gastrointestinal tract.
CC -!- INDUCTION: Activated by hedgehog; repressed by itself (Probable).
CC -!- PTM: Glycosylation is necessary for SHH binding.
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; U40074; AAC59898.1; -; mRNA.
DR PIR; T18538; T18538.
DR Ensembl; ENSGALG00000012620; Gallus gallus.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.

```


Db 321 SKYVHMQEELII---GGTVKNASQGI-----VSALALQTMFQLWTP 359
Qy 550 VFPFLAIGYKGYKDYSEAEALIMTFSNNYPAGDPLRLAQAKLWBEAFLEENRPAQRNAG 609
Db 360 KQMY-----EHFKGHEVVSHWNNE-----DKAAALAEAWQTYVQVHQSVPQNSS 406
Qy 610 MFQVTFATERSLEDEINRTTADLPIFATSYIVIFLYISLALGSSYSSRSRVWVDSKATLG 669
Db 407 QKVLFPFTT--TTLDLILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CAKSQGAVG 460
Qy 670 LGGVAVLGAWMAAGFFSYLGIRSSILVILQWPFVLVSVDADNIFIFVLEYQLRPRPG 729
Db 461 LAGVLLVALSVAAGLGLSLGIGSNAAATQVLPFLALGVGDDVFLLAHAFSETGQNK 520
Qy 730 EPREVHIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSLAVILDFLLQMS 789
Db 521 IPFEDRTGECLKRTGASVALTSISNTAFFMAALIPALRAFSLQAQAAVVVFNFAVLL 580
Qy 790 AFVALLSLDSKQEAASRLDVCCC-----VKPQ-----ELPP----- 820
Db 581 IPFALLSMDLYRREVRLDIFCCFSPCVSRVQIEPOAYTDNDNTKYSLPPTYSSHSP 640
Qy 821 -----PGQ----- 823
Db 641 AHETOITMQSTVOLRTEYDPRDTQLYYYTAQPRSEISVQPAASTQDVSGQTPESTSTRD 700
Qy 824 -----GEGLLGFFQKAYAPFLLHWTIRGVVLLFLLFALFGVSLYSMCH 866
Db 701 LISQFSVHGGSMQCTPDSKWTLSSPAETHYAPFLPKTKKAVILGLFALLSVSLYGTTR 760
Qy 867 ISVGLDQELALPKDSYLDYFLFNRYFEVGAUVYFVITLGYNFSSEAGMAICSSAGCN 926
Db 761 VRDGLDLTDIVPRETREYDFIATQKYPF-YHMYVVVTQKA----- 800
Qy 927 NFSFTQKQOYA--TEPFQESYLAI PASS-----WVDDFDIMLTPSS----- 965
Db 801 DYPAQRLLYELHKKFVGVYVLLSGNKQLPKWHLHYFRDMLQGLQDIFHEWEAGKITR 860
Qy 966 -----CCRLYISGPNKPCPSTVNSLNCNCSIMTGWSPRPSVEQPHKYL- 1012
Db 861 NDRNASDDAVLAYKLLIQTGNSDK--PINLNQLT--KQRLVDADGIIQPN--FYIYLT 914
Qy 1013 PWFLEDRPNIKCPKGLAASVSV-----VAIASQANIRPHPEWHLKADRPETRTIRAAEPYVQPP 964
Qy 1051 AYHKPLKNSQDYTEALRAARELAANITADLRVPGTDPAPFVFPYTTITNVFYEQLTILP 1110
Db 965 FYINGLRSETSDFEALKEKRAICNNVTS-----LGVSSYPNGYFP-----LPWEQVLSLRH 1015
Qy 1111 EGLFMLSCLVPTFFAVSCLLGLDLRSLGLLNLSIVMILVDTVGFMALWDISYNAVSLIN 1170
Db 1016 WLLLSISVVLACTFLVLCALFLNPNWTAGII--VMVLALMTVELFGMGLIGIKLSAPVVI 1074
Qy 1171 LYSACMSVEFVSHITRSP--AISTKPTWLERAKEATISGSAVFAVAMTNLPGILVLG 1228
Db 1075 LTASVIGIEFTVHVALAFLTAAGDKN---RRAVLALEHMFAPVLDG-AVSTLLGVMLA 1130
Qy 1229 LAKAQIQLIFFRNLNLLTLLGLLHGLFVLPVILSYVG--PDVNPA-----LALQKRAE 1281
Db 1131 GSEFDFIVRYFFAVLAILLGLVNLGLVLLPVLSSFFGYPYPEVSPITNGGDRLEVSQDPPP 1190
Qy 1282 EAVAAMVASCPNHPRSVST-ADNIYVNHSPFEGSING 1317
Db 1191 N-----IVRFPHEPRPTQNGSDSDSEYSSQTSVSG 1221

RESULT 68

Q98SW6_XENLA

ID Q98SW6_XENLA PRELIMINARY; PRT: 1418 AA.

AC Q98SW6;

DT 01-JUN-2001 (T-EMBLrel. 17, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Patched-1.
GN Name=Ptcl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21095181; PubMed=11165486; DOI=10.1016/S0925-4773(00)00526-8;
RA Koebnick K., Hollemann T., Pieler T.;
RT "Molecular cloning and expression analysis of the Hedgehog receptors
PT Xpc1 and Xsmo in Xenopus laevis.";
RL Mech. Dev. 100:303-308(2001).
DR EMBL; AF302765; AAK15463.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR InterPro; IPR00731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
SQ SEQUENCE 1418 AA; 158226 MW; 4FEA30C2D5031F54 CRC64;
Query Match 8.9%; Score 617.5; DB 2; Length 1418;
Best Local Similarity 21.6%; Pred. No. 2.5e-34;
Matches 280; Conservative 184; Mismatches 452; Indels 381; Gaps 43;
Qy 306 RVAPARDKSNVDPKGTSLSDKLSFSTHTLGLQFFQG-----W----- 344
Db 21 RVRRRRGRSRVAPPDHDYLRQPSYCDANFALQOISEKAIGRAKAPLWLAFFQRLFKL 80
Qy 345 GTWVASPLTILVLSVIPVVALAAGLVFTELTDPVELWSAPNSQARSEKAPHQHFQ-- 402
Db 81 GCYIQKCGKFLVVGLLIFGAFVGLRAANLETNVEELWVEVGVRSRELDYTRQKIGEE 140
Qy 403 -----PP-----PFTNQVI 411
Db 141 AMFNPQLMIOTFLEDGANVLTEALLQHLHSALEATKVQVVMYKNPKWLELCFKSGELI 200
Qy 412 LTAPNRSY---RYDSLIL-----GPKNFSGILDL---DLL-----LELL-ELQE 449
Db 201 TEAVVSGIISMTPTCLITPLDPCFWEKAKLQSGMAYLPQKDILOWTDFPLELLELKK 260
Qy 450 RLRLHQVMSPEAQR-----NISQIDICYAPLNPDNTSLYDC-----CINS 489
Db 261 GKLDHDIWEEMINKAEVGHGYMDRCLNPSDKNCPYAPNKNSTKPDVSVLSLGGCYGL 320
Qy 490 LLQVFPQNRITLLLTANOTLMQGTQSQVDWKHFLYCANAPLTFKDGTTALALSCMADYGNP 549
Db 321 SKKYMHMQBELII---GGTVKNASQGI-----VSALALQTMFQLWTP 359
Qy 550 VFPFLAIGYKGYKDYSEAEALIMTFSNNYPAGDPLRLAQAKLWBEAFLEENRPAQRNAG 609
Db 360 KQMY-----EHFKGHEVVSHWNNE-----DKAAALAEAWQTYVQVHQSVPQNSS 406
Qy 610 MFQVTFATERSLEDEINRTTADLPIFATSYIVIFLYISLALGSSYSSRSRVWVDSKATLG 669
Db 407 QKVLFPFTT--TTLDLILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CAKSQGAVG 460
Qy 670 LGGVAVLGAWMAAGFFSYLGIRSSILVILQWPFVLVSVDADNIFIFVLEYQLRPRPG 729
Db 461 LAGVLLVALSVAAGLGLSLGIGISFNAATQVLPFLALGVGDDVFLLAHAFSETGQNK 520
Qy 730 EPREVHIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSLAVILDFLLQMS 789
Db 521 IPFEDRTGECLKRTGASVALTSISNTAFFMAALIPALRAFSLQAQAAVVVFNFAVLL 580
Qy 790 AFVALLSLDSKQEAASRLDVCCC-----VKPQ-----ELPP----- 820

```

Db 581 IPFALLSMDLYRREVRLDIFCCFSSPCVSRVQIEPQAYTNDNDNTRYSLPPTYSHSF 640
QY 821 -----PGQ----- 823
Db 641 AHETQITMSTVQLRTEYDPTQLYTTTAQRSEISVQPAASTPDQVSGQTPESTSSTRD 700
QY 824 -----GSGLLGPFQKAYAPLLHWITRGVVLLFLALFGVSLYSMCH 866
Db 701 LISQSVHGGSMQCTPDSKWTLSFAEKHYAPFLKPKTKVAVILGFLALLSVLYGTTTR 760
QY 867 ISVGLDOELALPKDSVLLDYFLFLNRYFEVGPVVFVTTILGYNFSSEAGNNAICSSAGCN 926
Db 761 VRDGLDITVPRETREYDFIATQFKYFSF-YHMYVYTQKA----- 800
QY 927 NFSFTQKIOYA--TEPPEQSILAIPASS-----WDDDFIDMLTPSS----- 965
Db 801 DYPRAQLLYELHKSFGVYRVLLLEGKQLPKWMLHYFRDMLQGLQDTPDHEWEAGKITR 860
QY 966 -----CRLYISGPNKQPCPSTVNSLNCNCSITMGSVRSPVQFHKYL- 1012
Db 861 NDYRNASDDAVLAYKLLIQTGNSDK--PINLAQLT--KQRLVDADGIIQFNA--FYIYLT 914
QY 1013 PWFLENDRPNKCPKGLAAYSTSV-----NLTSQGVLASRFM 1050
Db 915 AWVSNDP-----VAYASQANIRPHPEWLHDKADRPETRTIRAAPPIEYVQPP 964
QY 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPDPAFEVFPYITITVVFYEQYLTILP 1110
Db 965 FYLNGLRSTDFEALKEKRAICNNYTS-----LGVSYPNGYPP-----LFEQVILSRH 1015
QY 1111 EGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLIVDTVGFPMALDIIYNVSLIN 1170
Db 1016 WLLLSISVVLACTFLVCALFLNPNFTAGII-VVMLALMTVELFGMGLIGIKLSAVPVVI 1074
QY 1171 LYSVGMSTVEFVSHITRSP--ALSTKPTWLERAKERTISMGSAPVAGVAMTLPGLILVG 1228
Db 1075 LIASVGIGVEFTHVALAFELTAVGDKN---RRAVLALHEMFAPVLDG-AVSTLLGVMLLA 1130
QY 1229 LAKAQLIQIFFFFRLNLLITLLGLHLGFLPVLPSVVG--PDVNPA-----LALQEKRAE 1281
Db 1131 GSEFDFIVRYFFAVLAILLLGVNLGLVLLPVLPSFFGYPEVSPNTNGDRLEVSQDPPP 1190
QY 1282 EAVAAVMVASCPNHPRSVST-ADNIYVNSHFECSKING 1317
Db 1191 N-----IVRFEHPRPTQNGSDSSDSSEYSSQTSVSG 1221

RESULT 69
Q9W6T6 BRARE
ID Q9W6T6 BRARE PRELIMINARY; PRT; 1243 AA.
AC Q9W6T6
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Patched-2 protein.
GN Name=ptc2; Synonyms=ptcl;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lewis K.E., Concordet J.P., Ingham P.W.;
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ007742; CAB39726.1; -; mRNA.
DR Ensembl; ENSDARG0000016404; Danio rerio.
DR ZFIN; ZDB-GENE-980526-196; ptc2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.

```

```

DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patched_m_recept.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
SQ SEQUENCE 1243 AA; 138548 MW; FABF459DB0C91371 CRC64;

Query Match
Best Local Similarity 22.1%; Pred. No. 2.8e-33;
Matches 273; Conservative 174; Mismatches 413; Indels 375; Gaps 45;

QY 356 LVLVIVPVVALAAGLVFTTELTPDVELMSAPNSQARSEKAFHQHPG-PFFRTNQVILTA 414
Db 91 LVVGLLIIFGAFVGLRAANLETDVEKLWVEVGRVNOELKYTRQKIGBAMFSPQLMIQT 150
QY 415 PNRSSRYDSSLGPKNPSGILDLLELLELLELRLHLOWMSPEAQRNISLQDICYAP 474
Db 151 PRQEG-----ANILTVBALKQHLDSAKASRVHYTM--YNRQWTLLEHLCYKS 195
QY 475 ---LNPDMT-----SLYDCCINSLLQYFQNNRTL-----LILLTANQTLMGQTS 514
Db 196 GELVTETNVVDQILEKLHPCLVITPLDCWEGAKLHSGTFYLPKPKPLQWTFNPDMPGFTA 255
QY 515 -----QVD-WKHPLY-----CAN-----APLT-----FKDGTALALS 541
Db 256 ELKMLKYQVDSWEEMLVKADVGQGYMNRPCLNADPCPLSAPNKNNTTGFVDVAPVLTGG 315
QY 542 CMADYCAPVFPF-----LAIGYKGYKDYSE---AEALIMTFSLINNYPADGPLAOAKL-- 591
Db 316 C---YGLSKKYMHOEELIVGGTKKNDGKLLSAQAFOTMFOLMT-----PKQMYEHLKG 367
QY 592 -----WEB-----AFLEEMRAFQRMAGMFQ-----VTFTAERSLDEBINT 628
Db 368 YDEVSHINWEDKAAAILL---AWQRYKSEAVQSVNVSSQKVLFTFT--TILEDILKXP 423
QY 629 TAEDLPIFATSYVIVFYISLALGSYSWSRVMVDKATLGGVAVILGAVMAAMGPPFS 688
Db 424 SDVS VIRIATSGYLLMLAYACLTM---LRWD---CAKSGQAVGLAGVLLVTLVSAAGLGLCS 478
QY 689 YLIGISSLVILQVDFVLVSLVGCADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAPSM 748
Db 479 LIGISFNAAATQVLPFLALGVGDVDFLLAHAFSTGQNKRIFFEDRTGCKLRTGASV 538
QY 749 LCLSEALCFFLGALPMPAVRTFALTSGLAVIDFLLOMSAFVALLSDSKQESRLD 808
Db 539 LTSISNVTAFMAALIPALRAFSLOAAVVVVVFAMVLLIFPAILSDHLTRDRRRFD 598
QY 809 VCCC-----VKPQEL-----PPPG----- 822
Db 599 IFCCFVSPCANRVIQLEPQAAVADSSADSSRYSPPPSYSSSHSFAQHTQITMQSTVQLRTE 658
QY 823 -----QGE----- 825
Db 659 YDPRQTQAYTTGEPHSHISVQPYAPNTNPNPNRNNNDNCTNSSSSAVPGVATDTASC 718
QY 826 -----GL-----LLGFQKAYAPFLHWTIRGVLVLLFLA 855
Db 719 QSPDGASSTRDLLSQFGSGIKCLSPYSRWTTFASAEKHYPALLOSTTKVVVIFLFLA 778
QY 856 LFGSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGPVVFVTTILGYNFSSEAG 915
Db 779 LAGVSLYGTTRVRDGLLELTDIVPRETGEYDFPRAQFVY----- 816
QY 916 MNAICSSAGCNPSFTQKIQYATEPPE-----QSYLAIP-----ASSWVDDFIW 960
Db 817 -----SSFYNNYVVTQVRVDVAIQPOLYELHQFSGVKYILIREENGQLPRMMPHYFRDW 870
QY 961 L-----TPSS-----CCRLYISGPNKQPCPSTVNSLNCNKM 994
Db 871 LIGLQEAERDKQWQAGRIITQGNTRNGTDDGVLAYKLLVQTRGREK-----TITRQL 921
QY 995 SITMGSVRSPVQFHKYL-PWFLND-----RPNIKCPKGLAAYSTSV-----N 1037

```

Db	922	VSADGIINPNA--FYIYLSAMVSNDFVAYAAASQANIRPH-----PWEHLHRTDTSIPASRLN	976
Qy	1038	LTSDQVQLASRFMAYHKPKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVFPYTI	1097
Db	977	IPAAEPIEYAQFPFVLNGLRTPQVEVAIESVRAICNNYSRQ-----GLPSYPNGY	1027
Qy	1098	TNVFVEQYLTLPGLFMLSCLVPTFAVSCILLGLDLRSGLLNLISVMILVDTVGEMA	1157
Db	1028	PFLEWEQYVGLRHMLLSISVLACTIONCAVFLINPWTAGII-VLVLSTMTVELFGMMG	1086
Qy	1158	LWDSYNVAVSLNLVSAGMSVEFVSHITRSPASTKPTWLERAKEATISMGSAVEAGV-	1216
Db	1087	LIGIKLSAVPVVILIASVIGVEFVTHVALAFLTAIG-----DRNKRAVLAL-EMHAPVL	1141
Qy	1217	--AMTNLPGIILVGLAKAQLQIIPFRNLNLITLGLHLGLVFLPVILSYVVG--PDVNP	1272
Db	1142	DCAPSTLGLVLMAGSEDFIVRYFFAVLAIIITVLGVNLGLVLLPVLLSYFGPCPEVSPA	1201
Qy	1273	---LALQKRAFEAAVAAVVASCPNHPSRVSTADN	1304
Db	1202	DGRSLPTSPBPQPPQVVRFTMRPSHTTPEGAGSDS	1236
RESULT 70			
Q9DEF3_XENLA PRELIMINARY; PRT; 1413 AA.			
AC	Q9DEF3		
DT	01-MAR-2001	(TremBLrel. 16, Created)	
DT	01-MAR-2001	(TremBLrel. 16, Last sequence update)	
DT	01-OCT-2003	(TremBLrel. 25, Last annotation update)	
DE	Patched-2.		
GN	Name=Xptch-2;		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
OC	Xenopodinae; Xenopus; Xenopus.		
OX	NCBI_TaxID=8355;		
NP	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=20500895; PubMed=1104611; DOI=10.1016/S0925-4773(00)00436-6;		
RA	Takahashi T., Takahashi T.C., Takabatake Y., Yamada K., Ogawa M.,		
RA	Takahashi K.;		
RT	"Distinct expression of two types of Xenopus Patched genes during		
RT	early embryogenesis and hindlimb development."		
RL	Mech. Dev. 98:99-104(2000).		
DR	EMBL; AB037688; BAB18575.1; -; mRNA.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008158; F:hedgehog receptor activity; IEA.		
DR	GO; GO:0004872; F:receptor activity; IEA.		
DR	InterPro; IPR004766; Patchedtm_recept.		
DR	InterPro; IPR000731; SSD 5TM.		
DR	Pfam; PF02460; Patched; 1.		
DR	TIGRFAMs; TIGR00918; ZA060602; 1.		
DR	PROSITE; PS0156; SSD; 1.		
SQ	SEQUENCE 1413 AA; 157372 MW; 1215D2D7733E7F17 CRC64;		
Query Match 8.7%; Score 599.5; DB 2; Length 1413;			
Best Local Similarity 22.0%; Pred. No. 4.6e-33;			
Matches 236; Conservative 180; Mismatches 390; Indels 347; Gaps 44;			
Qy	355	ILVSLVPVVALAAGLVTELTATDPVELMSAPNSQARSEKAFHQHFG-PFRFTNQVILT	413
Db	83	VLFIGLLVFGALAVGLRVASIEDIERLWVEAGSRVSHELRYTKELGESVVTYQMLIQ	142
Qy	414	APNRSSRYDSLLGLKPNFGSGLDLDLLELLELQERLHLQWSPQAQNTSLQDICY-	472
Db	143	TPKREG-----ENILTHEAL--LHLRAALAASKVQVSMYKSGMDLNKICYK	187
Qy	473	--APLNPDN-----TSLYDCCNLSLQVFNQNTL-----LLLTANQTL-----	509

RESULT 71
Q6IRA5_XENLA


```
RESULT 72
P91129 CABEL PRELIMINARY; PRT; 933 AA.
ID P91129 CABEL PRELIMINARY; PRT; 933 AA.
AC P91129;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Patched related family protein 2.
GN Name=ptr-2; ORFNames=C32E8.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; U88308; AAB42325.1; -; Genomic_DNA.
DR PIR; T25600.
DR Ensembl; C32E8.8; Caenorhabditis elegans.
DR WormBase; WBGene0004217; ptr-2.
DR WormPep; C32E8.8; C808532.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR001036; Acrlavin_res.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 933 AA; 104354 MW; C12E5376315FDD11 CRC64;

Query Match 8.7%; Score 598.5; DB 2; Length 933;
Best Local Similarity 22.9%; Pred. No. 3.1e-33;
Matches 251; Conservative 181; Mismatches 362; Indels 303; Gaps 37;

QY 341 FCGWGTWASWPLTLLVLSVPVVALAAGLVFTLTTPVELWSAPNSQARSEK-AFH-- 397
DB 14 FRLGLICDHPLPFPVPLPFTAAAGVGLHLNPLSDAVLFTPLGAGQKMERMSIHEK 73

QY 398 ----DQHPGP---FRTNQVILTAPNRSYRYSLLGPKNPSGILDLDLLELLEQER 450
DB 74 WPLTDNNYIPGRAVYQTSREIQVTALRN----DSNILDPKFANAVYQLDKY-----IQTR 124

QY 451 LRHLQWSPQAQRNLSLODICVAPLNPNSTSLYDCINSLLQYFONNRTLLLTANQTLM 510
DB 125 VRVLH----- 129

QY 511 GQTSQVDMKDHFLYCANAPLTKDGTALA-----LSCNADYG----APVPPFLAIGCV- 559
DB 130 -----NGHYYSYKNLCQYKNGCGPSNKHVHLSDLNHNHFNITYPYFRFGSEGGYI 181

QY 560 -----KQKDYSE-----AEALIMTFSLNNYP-----AGDPLRAQAKLWEAFLEE 599
DB 182 GSSLGGVTWKGENETDILASAKANFMVHLKFEEMSYISGE-----WE---LEL 230

QY 600 MFAQRMGMP-QVTFTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALSGY--SS 656
DB 231 GRMLTQYDPDVISITYFHSQTLADELKNADTLIPRFIISITLLIVFSTLCSLFIDGS 290

QY 657 WSRVMDSKATGLGCVAVVIGAVMAAGCFYSLGIRSLVLOVVPVLSVGNADNFI 716
DB 291 FSDVWLSKPLSIILGWSAGAITGTGVLGMLGMPYN-DIVGVMPLVLAVGVNDNML 349

QY 717 FVLEYQRLPRRPGPREVH--IGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTAL 774
```


DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 896 AA; 100393 MW; 3B6B90BE1FF12EE6 CRC64;

Query Match 8.6%; Score 597; DB 2; Length 896;
Best Local Similarity 23.3%; Pred. No. 3.8e-33;
Matches 232; Conservative 182; Mismatches 378; Indels 206; Gaps 36;

QY 339 QFFQCGTGWASWPLTILVLSVIVVVALAAGVFTLTTPVELMSAPNSQARSSE-KAFH 397
DB 14 KFFHSGYLLVSGYPFIFLSPILVTCFLATGHLGIEPLSDALYLTPTDAPSKERRIIH 73

QY 398 DQHFQGFRTNQVILTAPNRSRYDLSLLGPKNFSGILDLLELLELQERLRLQVW 457
DB 74 DA--WPLVDGTFVAGRAVQTQS--REVQVAVVARSNGNILLDRVPSNELKLMESFIRN-NIT 128

QY 458 SPEAQRNLSLDICVAPLNPNTSLYDCCINSLQYFQNNRTLLLLLTANQTLMGQTSQVD 517
DB 129 VQFSNRTWSFADLCIA--GPDGR-----CANNDHIQ-----LASRLHQ-- 164

QY 518 WKDFLYCANAPLTTFKQGTALALSCMADYGAIVFPFLAIGGVK-----GKD-YSEAEAL 570
DB 165 ---HGINTYPTVRLSDKSAYIAS-----ALGGVKLAGNGENIIVEATAM 208

QY 571 INTFSLNNYP-----AGDPRLAQKLMEEAFLEEMRAFQRMAGFQVFTTFAERSLEDE 624
DB 209 LLIYQLKFPNEISYVSG-----LMEREPKMDXY-KKQAKYISITYFHSQTLSD 259

QY 625 INRTAEDLPATSYIVIFLY-----ISLALGYSNSRVMVDSKATLGLGGVAVVIG 678
DB 260 LNRNAERLAPKFIGAFVILVCFVSLCVISVITKSGYIDW-----VVTKPILSLVGSNAGM 315

QY 679 AVMAWAGPFSYLGIRSSLVILQVPLSVGADNIFIVLEYQRLPRPGEPREHICR 738
DB 316 GIASANGMLTYLEIQYN-DITAVMPFLVAVGTDNMLVMSLAKRTDR--NLKYDQRIAE 372

RESULT 74
P91346 CAEEL
ID P91346 CAEEL PRELIMINARY; PRT; 900 AA.
AC P91346;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Patched related family protein 10.
GN Name=ptr-10; ORFNames=F55F8.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Feloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; U00447; AAB37812.3; -; Genomic_DNA.
DR PIR; T29590; T29590.
DR Ensembl; F55F8.1; Caenorhabditis elegans.
DR WormBase; WBGene0004224; F55F8.1.
DR WormPep; F55F8.1; C36938.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 900 AA; 100415 MW; 12A301D83326A464 CRC64;

Query Match 8.6%; Score 592; DB 2; Length 900;

Best Local Similarity 23.3%; Pred. No. 8.6e-33;

Matches 237; Conservative 170; Mismatches 369; Indels 242; Gaps 36;

QY 739 ALGRVAPSMCLSLSEACFFLGLTALTPMPAVRTALTSGLAVIDFLQMSAFVALLSLD 798
Db 373 CMADAIVSILITDALDSFGVGTITTPAQVICIYTMCAILLTFAQLTFFCAILVY 432
QY 799 SKROE-----ASRLDV---CCVKPQELPPPGOG----- 824
Db 433 TRIBEQGLHSIWLRPVAVTSSTPLNVKLFWLGSQPOK-PLPSCGTVSSTSSVSTMTSQA 491
QY 825 -----EGLLGFFPKAKAPFLLH-WITRGVLLLFALFGVLSYMSCHISVGLDOE 874
Db 492 TSPASKHLHCAATSFRRWYAPVLMQPMI-RAIAGLWYLYLIGISYVGCSTHLKEGLEPA 550
QY 875 LALPKDSVLLDYELFLNR-YFEVGGAPVVFVTT-----LGNFSEAGNAICSSA- 923
Db 551 NLLVDDSYATPHRYVLEKHYHAGSLQIVVSNPPDLRDPVERINMDKMASTFANCKVAI 610
QY 924 GGNFSPF-----TKIQYATE-----FPEQSYLAIPASSWDDFDWLTPSSCCRL 969
Db 611 GDDSVQFWLREMOVSEETHKIQYDNEKFYDHAQAQYIYSDMQPWVWV----- 659
QY 970 YISGENKOKFCPSTVNSLNCNKMCSIITWGSVRPSVEQFKYLPWFLNDRPNKCPKGL 1029
Db 660 -----WGRNN----- 664
QY 1030 AAYSTSVNLTSGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDP 1089
Db 665 -----NSRIKTRFMICMRDITTKYQTEATFRIASRF-----EQ 704
QY 1090 FEVFPYTTITNVEYQYLTLPGLFMLSCLVPTFAVSCLLGLDLRSLGLNLSIVMIL 1149
Db 705 YNVVTYMLPLWLTQDQALVVPNTQDIIVAVACMLVISALLIPQVCSFWV-AVTIGSID 763
QY 1150 VDTGFMALWDSYNAVSLINLVSAGVMSVEFVSHITRSPALSTKPTWLERAKEATISMG 1209
Db 764 LGVLGFMILWNNYLDNAISMITTIMSVGFSDVYSAHITVAYVSKESTTSARVCDALGDLG 823
QY 1210 SAVFAGVAMTNLPGLVLGLAKAQIQTFFFRNLNLTLLGLLHGLVFLPVILS-YVG 1266
Db 824 NPVAQG-AMSTILAVSLSDVPAYMI-VTFFTKTVFLAISIGLHGLVFLPLMLSVFG 879
RESULT 75
Q4H343 CIOIN
ID Q4H343 CIOIN PRELIMINARY; PRT; 278 AA.
AC Q4H343;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nienann-Pick disease, type C3 (Fragment).
GN Name=Ci-Niemann-Pick Type C3;
OS Clona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15269171; DOI=10.1242/dev.01270;
RA Imai K.S., Hino K., Yagi K., Satoh N., Satou Y.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis."
RL Development 131:4047-4058(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12736827; DOI=10.1007/s00427-003-0330-z;
RA Satou Y., Satoh N.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis."
RL Dev. Genes Evol. 213:211-212(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Satou Y.;
RT "Expressed genes in Ciona intestinalis."
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB210579; BAE06584.1; -, mRNA.
FT NON TER 1
SQ SEQUENCE 278 AA; 30653 MW; 59APAG65596844232 CRC64;
Query Match 8.5%; Score 589; DB 2; Length 278;
Best Local Similarity 47.2%; Pred. No. 3e-33;
Matches 120; Conservative 42; Mismatches 76; Indels 16; Gaps 6;
QY 869 VGLDQELALPKDSVLLDYFLFPLFVGVGAPVVFVTTLGYNFSSEAGNAICSSAGCNF 928
Db 2 IGLDQSLMESDESVLYDFPGMNNYLSVGAPVVFVVDGQNYTDAAGANQICGGMGCNN 61
QY 929 SFTQKIQYATEFPEQSYLAIPASSWDDFDWLTP-SSCRLYISGNKDKFCPSTVNSL 987
Db 62 SLIEQIARMSKMPNYSYHIAVPASSWLDYDFWLPKQSSCCRHDTG-EEDVFCNATVST 120
QY 988 NCLKNMCSI--TMGSVRPSVEQFKYLPWFLNDRPNKCPKGLAAYSTSVNLTSDG--- 1042
Db 121 SCIA-CRSAQESANQSRPTDPFMKFLPWLNDNPETKCAKGHAAYGTVSKVIDEGKKS 179
QY 1043 QVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVFPYTTINVFY 1102
Db 180 RVGATSMAYHTLTSTKDFIGCLRSANKIAEIS-----QNTTAEVFPYSVVFY 231
QY 1103 EQYLTLPGLFML 1116
Db 232 EQYLTIVHDTIFNL 245
RESULT 76
Q61EN1 CAEBR
ID Q61EN1 CAEBR PRELIMINARY; PRT; 933 AA.
AC Q61EN1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG12004.
GN Name=CBG12004;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAC01000059; CAB66666.1; -, Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008159; F:hedgehog receptor activity; IEA.
DR GO; GO:0007275; F:development; IEA.
DR InterPro; IPR001036; Acrlivin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR PROSITE; PS50156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 933 AA; 104692 MW; 36364C782DC876DA CRC64;
Query Match 8.5%; Score 585; DB 2; Length 933;
Best Local Similarity 22.9%; Pred. No. 2.8e-32;
Matches 245; Conservative 186; Mismatches 370; Indels 268; Gaps 38;
QY 341 FQCGWGTWASWPLTILVLSVIPVVALAAGLVFTELTDDVELWASPNASQARSEK-AFH-- 397
Db 14 FRQLGFFVCDHPLFPIVFPFLFTTAMGVGLLHNLPLSDAVLFTPTGAQSKMERMSIHEK 73
QY 398 -----DOHFGP-----PFRNQVILTAPNRSRYRVDLLGPKNFSGILDLLLELLELQER 450

Db 74 WPLTNNYIPGRAVTSQRIQVTLARN-----DSNILDTPKANAVYQLDKY-----IQTR 124
QY 451 LRHLQWSPQAQNTSLQDICYAPLNPDPNTSLDYDCINSLLOYFQN-----NRTLILLTA 505
Db 125 VRVLNHHYYSYKNLCLO-----YKNGGCPSKNKHVHILSD 159
QY 506 NQTLMGQTSQVMDKHDFLYCANAPLTFKDTALALSCMADYCAPVFPFLAIGY-----559
Db 160 -----LYNHGFNTY-----PYRFGSEGYTGSSIG 186
QY 560 -----KGQDYS-----AEALIMTFSLNNYP-----AGDPRLAQAKLWEEAELEEMRAFO 604
Db 187 GVTVMKGENETDILASAKAWFIYHLKTHPEMSYISGEWELEGLMKD-YPED-----240
QY 605 RRMAGMFQVTFABRSLEDEINRTAEDLPFATSYVIFLYISIAL-----GSYS-SWS 658
Db 241 ----PYISITYFHSQTLADELKRNDTLVPRFVIGFTLLVWFSTLCSLCFIDGFSFIDW- 295
QY 659 RVWVDSKATLGLGGVAVVLGVAWMAAGPFSYLGIRSSVLVQLVPLVSVGADNIFPV 718
Db 296 ---VLSKPTLSILGVVNAIGAILTGICLSLIGMPYN-DIVGMPFLVAVGTDMFLMV 351
QY 719 LEYQRLPRRPGPREVH--IGRALGRVAPSMLLCSLSRAICFLGALTMPMPAVRTFALTS 776
Db 352 RAV-----RTSTHTVHERMGECMAADAVALITSTDLVLSFGVGTITIPAVQIFCVT 407
QY 777 GLAVILDFLQMSAFVALLSLDSKRQEARLSDVCCVK-----P 815
Db 408 GVAIFPAFLYQITFAACALAMKHEAGNSMPFVSPAEKRSLSLTFQIFNMGVSP 467
QY 816 QELPPPGQEGGLLGFQKAYAPFLHWTIRGVLLLP-LALFGVLSYMSCHISVGLDQE 874
Db 468 DHSSNDNVKQPTSRFFGEWYAPVLMPPFVRGIAWVFIYLLGAS-YGCSRIKEGLEPVP 526
QY 875 LALPKDSYLLDYFLFNRYF-EVGAPVYFV-----TTLGYNSSER- 914
Db 527 NLLVEDSTAIPHRLLEKFKWYGOQVQVIVNAPDLNRHSTRDRVHAMVLDFAFSKRAI 586
QY 915 GMAICSSAGCNGNFSTQKIQYATFEPQSYLAIPASSWVDVDFIDWLPSSCCRLYISGP 974
Db 587 GMSV-----QFWLFEMERYYQKELQV-----IID-----612
QY 975 NKDFCPTSVNLSLCKNCSITMGSVRPSVQFHKYLPWFILNDRPNLKCPKGGLA--AY 1032
Db 613 -----SSFYGLLQHFLASKTN-----NPLAEDIY 636
QY 1033 STSVNLTSDGVLAS-RFMAVHKPLKNSQDYTEALRAARELANITADLRKVGCTDPAFE 1091
Db 637 WGPMPDDNGTWQSFRTVGMKDLVTMDQTDATMSFRDVAARW-----PBFN 685
QY 1092 VPPYITITNVFYEQYITLPEGLFMLSCLIVPFAVSCLLGLDLRSGLNLISIVMLVD 1151
Db 686 VTFMPIMWFTDQYIIIPNTVQNIILALVMIVAVLFIPQPCS-LWVALACASIDFG 744
QY 1152 TVGFALNDISVNAVSLINLSAVGMSVEFVSHITRSPASTKPTWLERAKEATISMGSA 1211
Db 745 VIGYMTLGVNLDAISMITIIMSIGFSDVSAHAYGVVSTRDTATCKRQVQALSALGWP 804
QY 1212 VPAGVAMNLPGLVGLKAKAQLQIIPFRLNLLITLGLHGLVFLVILSYV--GPDV 1269
Db 805 LFQG-AMSTITAVSLADIPATMY-VTFPKTVLSISGLGLVFLVPLVLSIFVRGCC 862
QY 1270 NPALAEQKRAEAAVAAVWVASCNPNHPSKRVSTADMIYVN-----HSFE 1312
Db 863 IPSPPHHSAQKVEKQMKIAITSSPLHLQTVAPIRASSPISFPHRFE 911
RESULT 77
PTC2 HUMAN
ID PTC2 HUMAN STANDARD; PRT: 1203 AA.
AC Q9Y6C5; Q95341; Q95856; Q5QP87; Q6UX14;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Patched protein homolog 2 (PTC2).
GN Name=PTCH2; ORFNames=UNQ560/PRO1121;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
RX MEDLINE=99030620; PubMed=9811851; DOI=10.1073/pnas.95.23.13630;
RA Carpenter D., Stone D.M., Brush J., Ryan A., Armanini M., Frantz G.,
Rosenthal A., de Sauvage F.J.;
RT "Characterization of two patched receptors for the vertebrate hedgehog
protein family.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13630-13634 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND INVOLVEMENT IN MDB AND
BCC.
RX MEDLINE=99135908; PubMed=9931336; DOI=10.1093/hmg/8.2.291;
RA Smyth I., Narang M.A., Evans T., Heilmann C., Nakamura Y.,
Chenevix-Trench G., Pietsch T., Wicking C., Wainwright B.J.;
RT "Isolation and characterization of human patched 2 (PTCH2), a putative
tumour suppressor gene in basal cell carcinoma and medulloblastoma on
chromosome 1p32.";
RL Hum. Mol. Genet. 8:291-297 (1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2).
RX MEDLINE=99151521; PubMed=10029063;
RA Zaphiropoulos P.G., Uuden A.B., Rahman F., Hollingsworth R.E.,
Toftgard R.;
RT "PTCH2, a novel human patched gene, undergoing alternative splicing
and up-regulated in basal cell carcinomas.";
RL Cancer Res. 59:787-792 (1999).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
Vandlen R.L., Watanabe C., Wleand D., Woods K., Xie M.-H.,
Yanura D.G., Yu S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
effort to identify novel human secreted and transmembrane proteins: a
bioinformatics assessment.";
RL Genome Res. 13:2265-2270 (2003).
RN [5]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS GLN-346; LYS-493;
TYR-622; MET-988; MET-1019 AND MET-1121.
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT "NIHGS-SNP, environmental genome project, NIHGS RS15478, Department
of Genome Sciences, Seattle, WA (URL: http://esp.gs.washington.edu).";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG Human chromosome 1 international sequencing consortium;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May have a role in epidermal development. May act as a
receptor for sonic hedgehog (SHH).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9Y6C5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y6C5-2; Sequence=VSP_004542;
CC -!- DISEASE: Defects in PTCH2 are a cause of medulloblastoma (MDB)

[MIM:155255]. MDB is a malignant, invasive embryonal tumor of the cerebellum with a preferential manifestation in children. Although the majority of medulloblastomas occur sporadically, some manifest within familial cancer syndromes such as Turcot syndrome and basal cell nevus syndrome (Gorlin syndrome).

-1- DISEASE: Defects in PTCH2 are a cause of sporadic basal cell carcinoma (BCC) [MIM:605462].

-1- SIMILARITY: Belongs to the patched family.

-1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AF091501; AAC79847.1; -; mRNA.

EMBL; AF087651; AAD25953.1; -; mRNA.

EMBL; AF119549; AAD17260.1; -; mRNA.

EMBL; AY358555; AAQ88919.1; -; mRNA.

EMBL; AY438664; AAR05447.1; -; Genomic DNA.

EMBL; AL136380; CAI23127.1; -; Genomic DNA.

EMBL; AL592166; CAI23127.1; JOINED; Genomic DNA.

EMBL; AL592166; CAI13000.1; -; Genomic DNA.

EMBL; AL136380; CAI13000.1; JOINED; Genomic DNA.

Ensembl; ENSG00000117425; Homo sapiens.

DR HGNC; HGNC:9586; PTCH2.

DR MIM; 603673; -.

DR MIM; 155255; -.

DR MIM; 605462; -.

DR GO; GO:0005887; C:integral to plasma membrane; TAS.

DR GO; GO:0004888; F:transmembrane receptor activity; TAS.

DR GO; GO:0008544; P:epidermis development; TAS.

DR GO; GO:0006461; P:protein complex assembly; TAS.

DR GO; GO:0007283; P:spermatogenesis; TAS.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR004766; Patchedtm_recept.

DR InterPro; IPR000731; SSD_5TM.

DR Pfam; PF02460; Patched; I.

DR TIGRFAMs; TIGR00918; 2A060602; 1.

DR PROSITE; PS50156; SSD; 1.

KW Alternative splicing; Glycoprotein; Polymorphism; Receptor;

KW Transmembrane.

FT TOPO_DOM 1 57 Cytoplasmic (Potential).

FT TRANSMEM 58 78 Potential.

FT TOPO_DOM 79 392 Extracellular (Potential).

FT TRANSMEM 393 413 Potential.

FT TOPO_DOM 414 428 Cytoplasmic (Potential).

FT TRANSMEM 429 449 Potential.

FT TOPO_DOM 450 457 Extracellular (Potential).

FT TRANSMEM 458 478 Potential.

FT TOPO_DOM 479 501 Cytoplasmic (Potential).

FT TRANSMEM 502 522 Potential.

FT TOPO_DOM 523 531 Extracellular (Potential).

FT TRANSMEM 532 552 Potential.

FT TOPO_DOM 553 586 Cytoplasmic (Potential).

FT TRANSMEM 587 707 Potential.

FT TOPO_DOM 708 963 Extracellular (Potential).

FT TRANSMEM 964 984 Potential.

FT TOPO_DOM 985 991 Cytoplasmic (Potential).

FT TRANSMEM 992 1012 Potential.

FT TOPO_DOM 1013 1013 Extracellular (Potential).

FT TRANSMEM 1014 1034 Potential.

FT TOPO_DOM 1035 1064 Cytoplasmic (Potential).

FT TRANSMEM 1065 1085 Potential.

FT TOPO_DOM 1086 1093 Extracellular (Potential).

FT TRANSMEM 1094 1114 Potential.

FT TOPO_DOM 1115 1203 Cytoplasmic (Potential).

FT DOMAIN 394 552 SSD.

FT CARBOHYD 370 370 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 812 812 N-linked (GlcNAc...) (Potential).

FT VARSPPLIC 1143 1203 WGASSSLPQSFARVTTSMVAIHPPPLGAYTHPADPEPPW SPAATSSGNLSRRGPGPATG -> PEEI (in isoform

FT 346 346 VARIANT /FTId=VSP_004542.

FT 346 346 VARIANT E -> Q (in dbSNP:11573578).

FT 493 493 VARIANT /FTId=VAR_018935

FT 493 493 VARIANT E -> K (in dbSNP:11573581).

FT 622 622 VARIANT /FTId=VAR_018936.

FT 622 622 VARIANT H -> Y (in dbSNP:11573586).

FT 988 988 VARIANT /FTId=VAR_018937.

FT 988 988 VARIANT T -> M (in dbSNP:11573590).

FT 1019 1019 VARIANT /FTId=VAR_018938

FT 1019 1019 VARIANT V -> M (in dbSNP:11573591).

FT 1121 1121 VARIANT /FTId=VAR_018939.

FT 1121 1121 VARIANT I -> M (in dbSNP:11573598).

FT 2 2 CONFLICT /FTId=VAR_018940.

FT 120 120 CONFLICT T -> N (in Ref. 2).

FT 175 175 CONFLICT R -> L (in Ref. 2).

FT 753 753 CONFLICT R -> W (in Ref. 1 and 4).

FT 787 787 CONFLICT Q -> R (in Ref. 2).

FT 833 833 CONFLICT N -> S (in Ref. 2).

FT 837 837 CONFLICT P -> L (in Ref. 3).

FT 846 846 CONFLICT S -> G (in Ref. 2).

FT 897 897 CONFLICT D -> E (in Ref. 2).

FT 1203 1203 CONFLICT L -> F (in Ref. 3).

FT 1203 1203 CONFLICT G -> GDYKDDDK (in Ref. 4).

SQ SEQUENCE 1203 AA; 130544 MW; 4FAB06999782C031 CRC64;

Query Match 8.5%; Score 584; DB 1; Length 1203;

Best Local Similarity 22.9%; Pred. No. 4.7e-32;

Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;

QY 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLGQFFQG----WGTWVASWPLTILVLS 359

DB 3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLRAVPGQLLFSLGGCIORHCKGVFLG 62

QY 360 VIFVVALAAGLVFTLTDVPELMSAPNSQARSEKAFHDQHFQ--PFERTNQVILTAPNR 417

DB 63 LLAFGALALGLRMAIIEIENLEQWLWVEGVSQELHYTKELKGBEAAVTSQMLQTARQE 122

QY 418 SSVRYDSLILGPKNFGSLDLDLLELELERLHLQVWSPQARNISLQDICY---AP 474

DB 123 GEN-----ILTPEALG-----LHQAALTASKVQVSVLYGKSWDLNKICYKSGVP 166

QY 475 LNPDN-----TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520

DB 167 LIENGMIERMTKLPCCVILTPDCFWEGAK--LQGGSAYLPGRPDIQWNLDPQLLE 223

QY 521 -----HFLYCANAPLTPKDGATALALSCMADYGA 548

DB 224 ELGPPASLEGFPRELLDKAQGVQVYVGRPCILHPDDLHCPPSAPNHHSRQAPNVVAHLSGCG 283

QY 549 PVFPP-----LAIGYKGDYSE---AEALIMTF-----SLNNTYPAGDPRLAQA 591

DB 284 HGFSKFMHWQEBELLGGWARDPQGLLRALAQSTFLLMSPRQLYEHFRGDIYQTHDIGW 343

QY 592 WEEAFLEEMRAPQRMAGMFQVTFARSLR-DEINRTAEDLPFATS-----Y 640

DB 344 SEEQASTVLQAWQRRFVOLAQEALENASQQTHAFSSTLDDI-LHAFSEVSAARVWGY 402

QY 641 IVIFLYISIALGYSYSSRWVWDSKATLGLGVAVVLGAVMAAGFFSVLRSRLVLIQ 700

DB 403 LLMLAYACVTM---LRWD--CAQSGSVGLAGLVILVALAVASGLGLCALLGITFNAATQ 457

QY 701 VVPFLVLSVGADNIFVLEYQRLPRRPFGEPEVHIGRALGRVAPSMMLCSISEAICFPFL 760

DB 458 VLPFLALGIGVDDVFL--LAHAFTREALPGTPLQERMGECIQTGTSVLTSINNMAAF 515

QY 761 GALTWPVAVRTPALTSGLAVIDLFLQMSAFVALLSLDSKQEAERLDVCCC----- 812

DB 516 AALVETPALRAFSLOAAIIVWGCTFVAVMLVFPAILSLDLRRRCQLDLVLCFSSPCSAQ 575

QY 813 ----VKPQEL-----PPCQGGSEGL-- 827

DB 576 VIQILPQELGDDGTVPGVIAHLTATVQAFTHCEASSQHVVTILPPQAHLVPPSPDPLGSEL 635

QY 1103 BOYLTLPEGLFMLSCL--LVPTAVSCLLLGLDLRSLGLNLISVIMLVDTVGFMAWD 1160
 DB 955 BOYLGL--RRCLLAVCLLLVCTFLVCAALLNPNWTAGLI-VLVLAAMTVLFLGIMGFLG 1011
 QY 1161 ISYNAVSLINLSVAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSVAFAGVAMTN 1220
 DB 1012 IKLSAIPVILVASGIGVETVHVAGLF-LTQGSNLRRAHALEHTFAPVTDG-AIST 1069
 QY 1221 LPGILVLGLAKAQLIQIPFFRLNLITLLGLLHGLVFLPVLVSYPVDPNPALEQKRA 1280
 DB 1070 LGLGLMLAGSHDFDFTVRYFFAALTVLLTLLGLHGLVLLPVLISILGPP--DEVIQMYKES 1127
 QY 1281 EEAANA--NVAACPNNHPSRVSTADNIYVN 1308
 DB 1128 PEILSPAPPOGGGURMGASSLSPOSFAFVTTSMTVIAIH 1165

RESULT 79

PTC2_MOUSE STANDARD; PRT; 1182 AA.

AC O35595; Q54670;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Patched protein homolog 2 (PTC2).
 GN Name=PTC2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=98122566; PubMed=9462734;
 RA Motoyama J., Takabatake T., Takekuma K., Hui C.-C.;
 RT "Ptc2, a second mouse Patched gene is co-expressed with Sonic
 hedgehog.";
 RL Nat. Genet. 18:104-106(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=129/SvJ; TISSUE=Liver;
 RX PubMed=12438747; DOI=10.1159/000064064;
 RA Frohlich L., Zhanquin L., Beier D.R., Lanske B.;
 RT "Genomic structure and refined chromosomal localization of the mouse
 Ptc2 gene.";
 RL Cytogenet. Genome Res. 97:106-110(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 196-446.
 RC STRAIN=BA1B/c; TISSUE=Neuroretina;
 RX MEDLINE=97379366; PubMed=9237688; DOI=10.1016/S0014-5793(97)00645-5;
 RA Takabatake T., Ogawa M., Takahashi T.C., Mizuno M., Okamoto M.,
 RA Takeshima K.;
 RT "Hedgehog and patched gene expression in adult ocular tissues.";
 RL FEBS Lett. 410:485-489(1997).
 CC -|- FUNCTION: May have a role in epidermal development. May act as a
 receptor for Sonic hedgehog (SHH).
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- TISSUE SPECIFICITY: Expressed in epithelial cells of the
 developing hair, tooth and whisker.
 CC -|- DEVELOPMENTAL STAGE: Detected in 8.5 to 17.5 dpc embryos.
 CC -|- SIMILARITY: Belongs to the patched family.
 CC -|- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

DR EMBL; AJ133482; CAC88120.1; -; Genomic DNA.
 DR EMBL; AJ133483; CAC88120.1; JOINED; Genomic_DNA.

DR EMBL; AJ133484; CAC88120.1; JOINED; Genomic DNA.
 DR EMBL; AJ133485; CAC88120.1; JOINED; Genomic_DNA.
 DR EMBL; AB010833; BAA24691.1; -; mRNA.
 DR PIR; I13952; I13952.
 DR Ensembl; ENSMUSG00000028681; Mus musculus.
 DR MGI; MGI:1095405; Ptc2.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR004766; Patchedtm_recept.
 DR InterPro; IPR000731; SSD_5TM.
 DR Pfam; PF02460; Patched; 1.
 DR TIGRFAMs; TIGR00918; 2A060602; 1.
 DR PROSITE; PS0156; SSD; 1.
 KW Glycoprotein; Receptor; Transmembrane.
 FT TOPO_DOM 1 57 Cytoplasmic (Potential).
 FT TRANSMEM 58 78 Potential.
 FT TOPO_DOM 79 394 Extracellular (Potential).
 FT TRANSMEM 395 414 Potential.
 FT TOPO_DOM 415 428 Cytoplasmic (Potential).
 FT TRANSMEM 429 449 Potential.
 FT TOPO_DOM 450 457 Extracellular (Potential).
 FT TRANSMEM 458 478 Potential.
 FT TOPO_DOM 479 501 Cytoplasmic (Potential).
 FT TRANSMEM 502 522 Potential.
 FT TOPO_DOM 523 531 Extracellular (Potential).
 FT TRANSMEM 532 552 Potential.
 FT TOPO_DOM 553 686 Cytoplasmic (Potential).
 FT TRANSMEM 687 707 Potential.
 FT TOPO_DOM 708 963 Extracellular (Potential).
 FT TRANSMEM 964 984 Potential.
 FT TOPO_DOM 985 991 Cytoplasmic (Potential).
 FT TRANSMEM 992 1012 Potential.
 FT TOPO_DOM 1013 1013 Extracellular (Potential).
 FT TRANSMEM 1014 1034 Potential.
 FT TOPO_DOM 1035 1064 Cytoplasmic (Potential).
 FT TRANSMEM 1065 1085 Potential.
 FT TOPO_DOM 1086 1086 Extracellular (Potential).
 FT TRANSMEM 1087 1107 Potential.
 FT TOPO_DOM 1108 1182 Cytoplasmic (Potential).
 FT DOMAIN 394 552 SSD.
 FT CARBOHYD 370 370 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 812 812 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1182 AA; 128586 MW; 715233D912C352F2 CRC64;

Query Match Similarity 8.4%; Score 583.5; DB 1; Length 1182;
 Best Local Similarity 24.2%; Fred. No. 5e-32;
 Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SFSTHTLLGQ-----PFQ-----WGTVVASWPLTILVLSVPPVVALAAGLVFTLT 376
 DB 20 SSAPHILAGSLQAPLWLRAYFQGLLFLSGCRQKHCGKVLFLGLVAFGLALGLRVAVIE 79
 QY 377 TDPVELMSAPNSQARSEKAFHDQHFQ-PFFRTNQVILTPNRSYRYDLSLLGPKNFSGI 435
 DB 80 TDLEQLWVEGSRVSRVQELHYTKELGEEAAVTSQMLI-----QTAHQEGNVLTPE----A 131
 QY 436 LDLDLLELLELOERLHLQWSPQAQRNLSIDICY---APLNPDN-----TSLYDCC 486
 DB 132 LD-----LHLQALTASKQVSLYSGKSWDLNLCYKSGVPLTENGMIERMIKLFPCV 184
 QY 487 INSLQYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---FKDGTALA 539
 DB 195 ILTPLDCEWGAQ---LOGGSAYLPGRPDIQWNLDPQOLLEELGPFASLEGFRELLDKA 241
 QY 540 LSCMADYGAQVY-----PFLA---IGGYKGDYS----- 565
 DB 242 QVGQAYVGRPCLDPDHPCHPSPAPNRHSRQAPNVAQELSGGCHGFHFKFMHMQEELLGG 301
 QY 566 -----EAEALIMTF-----SLNNYPAGDPLRAQAKLWEEAFLEEMAFORFAMAG 609
 DB 302 TARDLQQLRAEALQSTFLLMSRPLQYEHFRGDYQTHDIGWSESEQASVVLQAWORRFVQ 361
 QY 610 MFQVTFTAERSIE-----DEINRTAEDLPIFATSYIV--IFLYISLALGSYSWS 658

Db	362	LAQBALPANASQQIHAFSSSTTLDLIDRAFSE-----VSTTRVVGYYLMLAYACVMTLWRWD	417
Qy	659	RVMTDSKATLGLGGVAVVLGAVMAAGFFSVLGIIRSLVLTQVVPFLVSLVSGADNIFIV	718
Db	418	--CAQSOQAVGLAGVLLVALAVASGLGICALLGITFNAATQVLPFLAIGVDDIIFLLA	475
Qy	719	LEYORLPRRPGEPREVIHGRALGRVAPSMILCSISEAICFFLGALGALTMPAVRTFALTSGL	778
Db	476	HAFTKAP--PDTPLPERMGECLRTGTGSTVALTSVNNVAFPMFMAALVIPALRAFSLQAAI	533
Qy	779	AVILDFELLQNSAFVALLSLDSKQESRLDVCCC-----VKPQE-----	817
Db	534	VVGCNFAAVMLVFPAILSLDLRRHRQRQLDVLCCFSFPCSAQVITQMLPQELGDRVAVPGI	593
Qy	818	-----LPP-----PG-----QGBG--	826
Db	594	AHLTATVQAFTHCEASSQHVVITLIPQOAHLLSPASDFLGSSELVSPGGSTRDILLSQEBGTG	653
Qy	827	-----LTLGFFQKAYAFPLHLMITRQVLLFLFALFVGLSVLSMCHISVGLDQ	873
Db	654	POAACRPILCAHWTLAHFARVQFAPLJJLQTRAKAVLLFFCALLGLSLYGATLVQDGLAL	713
Qy	874	ELALPKDSYLLDYFLFNRPFEVGAIFYVTTLGYNFS--SEAGNNAICSSAGCNFSFTQ	932
Db	714	TDVVPRTGKEHAFSLAQRLYFSL--YEVALVTQGGFDYAHSORAL-----FDLHQ	761
Qy	933	KIQVATEFPBQSYLAI PAS-----SWVDDF-----IDWLTPSSCCRLYISGPNK	976
Db	762	RFSSL-----KAVLPFPATQAPRTWLHYHYSWLGQIQAAFDQDMASGRITCHSYRNGSD	816
Qy	977	D-----KFCPSVNSLNCXNCMSITM-----GSVRPSVEQFHKYLPWFNLDRPNKCPK	1026
Db	817	GALAYKLLIQTGNAQEPIDPSOLTTRKLVNKEGLIPP--ELFYMLTVMVSSDPL-----	869
Qy	1027	GGALAA-----YST-----SVNLTSQGVLASRFMAYHKPLKNSQDYTEALR	1067
Db	870	--GLAASQANFYPPPEWMLHDKYDITGLENRAIPAAQPLEFAQFPFLLHGLQKTADFVEAIE	928
Qy	1068	AARELAANI--TADLRKVPGTDPAPFVPPYITVTFYEQYLTILPEGLFMLSIC--LVPTF	1124
Db	929	GARAACTEAQOAGVHAYPSGSPF-----LFWEQYLG--RRCFLLAVCILLVCTF	976
Qy	1125	AVSCLLGLDLRSGLMLLSIVMLIVDTVGFMAWDISYNAVSLINLVSAGVMSVRFVSH	1184
Db	977	LVCALLLSFPWTAGLI--VLVAMTVBELFGIMFGELGKSAIPVILVASIGIGVFTTVH	1035
Qy	1185	ITRSFAISTKPTWILERAKEATISMGSAVFAGV---AMTNI.PGILVILGLAKAQLIQIPFFR	1241
Db	1036	VALGFLTSHGSRNLRAA-----SALEQTFAPVTDGAVSTILGLMLAGSNFDFIIRYFFV	1090
Qy	1242	LNLLITLGLLHGLVFLPVILSVVGP	1267
Db	1091	VLVTLTLLGLHGLLFLVLSILGP	1116

```

RESULT 80
Q9XXR9 CAEEL
ID Q9XXR9 CAEEL PRELIMINARY; PRT; 1003 AA.
AC Q9XXR9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein ptr-19.
GN Name=ptr-19; ORFNames=Y39A1B.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;

```

RG	The C. elegans sequencing consortium;
RT	"Genome sequence of the nematode C. elegans: a platform for
RT	investigating biology."
RL	Science 282:2012-2018(1998).
DR	EWBL; AL021482; CAAL6339.1; -, Genomic_DNA.
DR	PIR; T26746; T26746.
DR	Ensembl; Y39A1B.2; Caenorhabditis elegans.
DR	WormBase; WBGene00004233; Y39A1B.2.
DR	WormPep; Y39A1B.2; CE19144.
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0008158; F:hedgohog receptor activity; IEA.
DR	InterPro; IPR003392; Patched.
DR	pfam; PF02460; Patched; 1.
DR	PROSITE; PS50156; SSD; 1.
KW	Complete proteome; Hypothetical protein.
SQ	SEQUENCE 1003 AA; 112257 MW; 92EAF1EALP7627C CRC64;
Query Match	8.3%; Score 573.5; DB 2; Length 1003;
Best Local Similarity	22.8%; Pred. No. 2e-31;
Matches 236; Conservative 199; Mismatches 376; Indels 225; Gaps 42;	
Qy	339 QFQCGWTG---VASWPLTLVLVSIVPVVALAAGLVFTE--LTTDPVELWSAENSQARSE 393
Db	9 RFAHAFGSGVVRHPFPFPFPIIPILITAGLSGLLRHEQAFMKDELSLFTPTDAQARKE 68
Qy	394 KA-----FHDQHGCPPERTNQ-----VILTAPNRSSVRYDSLGLGKPNPSGILDLDL 440
Db	69 LSQDLHLFHNDSDFPVATRYDIRAGYIIVT-----NQEEDGDILNPLVMSISAMQLWS 123
Qy	441 LLELLELQERLHLQVMSPEAQRNLSLDIQC-----YAPLNPDNTSLYDCCI 487
Db	124 IVQSLTVED-----EDDRRIYPSICVKFPPIPPPEFSKALHSLPAPNMTTPEEICV 173
Qy	488 -NLLQYFQNNRTLLLTANQTLMGQT-----SQVDKDHFLYCANAPLTFKDGFTALAL 540
Db	174 SNFLPIEFK-----LLIVSDRSFINSIDEMTLSQI-----SDAIFQDSGGWTHL 218
Qy	541 SCMADYGAPVFPFLATGGY--KGKDYSAEALIMFTSLNNYPAGDPRLLAQAKLWEAFL 597
Db	219 -----LQGVTLDDKKIAGKAKMLLPYALRH-----SSDEDWVAEKWEVRLA 261
Qy	598 EEMRAFQRRMAGMQVPTTAERSLEDEINRTTAEDLPFATSVIVFLYISLALGSYSSW 657
Db	262 DFLQYD--SPIIRASWWTYETLAESAARDRLQLIHMLLPCFVCVSIPI-TIACCVCFSW 317
Qy	658 SRVWDSKATLGGVAVVLGAVMAAGPFSYIGIRSSLVILQVDFVLVSUGADNIFIP 717
Db	318 RR----SRPWLATGGVISAAATASAVGILLAGYGMTSVAYS-MPFIYFVSVDVNVFIL 372
Qy	718 VLVSQRLPRPGPREVHIGRALGRVAPSMLLCSLEAICFFLGALTTPAVRTEALTSG 777
Db	373 LSAWRSTST--ETLEHRMKETPADAAVSTVTSLTDLISFGVGCACTPPSPVQMFCAYAV 430
Qy	778 LAVILDFLQMSAFVALLSDSKRQBSARLDVCCVC-----KPELPPPGQCGEG----- 826
Db	431 AAVIFYYIQLTFPAVMVYTNREINNRR-----HCIFPHKLKDTLPEKIAAQGDRSFEK 486
Qy	827 --LLLGFFQKAYAPFLHWTITRGVILLFLALFGVSLYSNMCHISVGLDQELAPKDSYLL 884
Db	487 NTLAQPFRTTYSDFLNLPLVIRVILVTFCVILGVASVGCYTKVGLERPNLLPENSYGK 546
Qy	885 DYFLPLNRYP-EVGAQPVYFTLIGYNFSS-BAGMNAICSSAGCANNFPTQKIQYATEFPE 942
Db	547 RTLMMAEKYPSDYGSSLH-----VMYMLNSVDVAPRKIWN-----VLEKVELYE 592
Qy	943 QSYLAIPASSWDDFDIW-----LTPSSCCRLYISGPNKQFCPSTVNSLNLCLNCMS 995
Db	593 HTEFTASSDSWLTATFLAVKQAGLLITPE-----NFVYLKQ--- 629
Qy	996 ITWGSVRPSVEQPHKPLPWLNDRPNIKCPKGLAAYSTVNLTSQGVL-ASRFWAYHK 1054
Db	630 -----VELSOPQ-----AKYNRDVLVTGEGHEASRI----- 658

```
QY 1055 PLK-----NSQDYTEALRAARELANITADLRKVPCTDPAFVFPVYITNVTVEQKLTILP 1110
Db 659 PVQLRHVGSANOSRAMRLFRRLAE--TSBLQ-----TGVYADFFQ-----FAEQYNVLP 706
QY 1111 EGLFMLSCLVPTFAVSCILGLDLRSGLNLLSVTLLVDTVVGPMALWDLSYNVSLIN 1170
Db 707 GTLSIAVAGVAVVSLILIEPVAS-LWVSFVSIVSINIGLGFMTTWSVRLDFISWVT 765
QY 1171 LVSAVGMSEVFVSHITRSPASTKPTWLERAKEATISMSGSAVFAGVATNLPGLILVGLA 1230
Db 766 IVMSGFCVDFAAHLAYNFAKQNGDSRMRNALYAVGAPILMS-AUSTIIGVSFMSA 824
QY 1231 KAQLIQIIFERLNLITLILGLHGLVFLPVLS--YVGDVNPALALQKRAEBAVAV- 1287
Db 825 ESYVFR-SPLKTMILVILGALHGLVILPVLVLSMFYCGG-----SSKKAKEHIDAVD 875
QY 1288 -MVASCPNHPRSVST 1301
Db 876 QKLQAYNNPARTAS 890

RESULT 81
PTC1 BRARE
ID PTC1 BRARE STANDARD; PRT; 1220 AA.
AC Q98864
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Patched protein homolog 1 (Patched 1) (PTC1).
GN Name=ptc1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=96379744; PubMed=8787757;
RA Concordet J.-P., Lewis K.E., Moore J.W., Goodrich L.V., Johnson R.L.,
RA Scott M.P., Ingham P.W.;
RT "Spatial regulation of the zebrafish patched homologue reflects the
RT roles of sonic hedgehog and protein kinase A in neural tube and somit
RT patterning."
RL Development 122:2835-2846(1996).
CC -!- FUNCTION: Acts as a receptor for sonic hedgehog (SHH), indian
CC hedgehog (IHH) and desert hedgehog (DHH). Associates with the
CC smoothened protein (SMO) to transduce the hedgehog's proteins
CC signal (By similarity).
CC -!- SURCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Detected in embryonic presomitic mesoderm,
CC neuroectoderm, tissue surrounding the notochord, ventral neural
CC tube.
CC -!- DEVELOPMENTAL STAGE: At all stages, expression corresponds to the
CC localization of SHH. First detected during gastrulation. By 36
CC hours, PTC1 appears in the first branchial arch and the posterior
CC mesenchyme of the fin bud; by 48 hours, in the hindbrain and
CC foregut.
CC -!- INDUCTION: Activated by Sonic hedgehog.
CC -!- PTM: Glycosylation is necessary for SHH binding (By similarity).
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X98883; CAAG7386.1; -; mRNA.
CC FIR; T18291; T18291.
CC DR Ensembl; ENSDARG00000011590; Danio rerio.
```

```
DR ZFIN; ZDB-GENE-980526-44; ptcl.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRfams; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
KW Glycoprotein; Receptor; Transmembrane.
FT TOPO_DOM 1 84 Cytoplasmic (Potential).
FT TRANSMEM 85 105 Potential.
FT TOPO_DOM 106 419 Extracellular (Potential).
FT TRANSMEM 420 440 Potential.
FT TOPO_DOM 441 449 Cytoplasmic (Potential).
FT TRANSMEM 450 470 Potential.
FT TOPO_DOM 471 484 Extracellular (Potential).
FT TRANSMEM 485 505 Potential.
FT TOPO_DOM 506 528 Cytoplasmic (Potential).
FT TRANSMEM 529 549 Potential.
FT TOPO_DOM 550 558 Extracellular (Potential).
FT TRANSMEM 559 579 Potential.
FT TOPO_DOM 580 739 Cytoplasmic (Potential).
FT TRANSMEM 740 760 Potential.
FT TOPO_DOM 761 1016 Extracellular (Potential).
FT TRANSMEM 1017 1037 Potential.
FT TOPO_DOM 1038 1044 Cytoplasmic (Potential).
FT TRANSMEM 1045 1065 Potential.
FT TOPO_DOM 1066 1072 Extracellular (Potential).
FT TRANSMEM 1073 1093 Potential.
FT TOPO_DOM 1094 1110 Cytoplasmic (Potential).
FT TRANSMEM 1111 1131 Potential.
FT TOPO_DOM 1132 1143 Extracellular (Potential).
FT TRANSMEM 1144 1164 Potential.
FT TOPO_DOM 1165 1220 Cytoplasmic (Potential).
FT DOMAIN 421 579 SSD.
FT COMPBIAS 624 695 Thr-rich.
FT CARBOHYD 397 397 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 865 865 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 888 888 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1220 AA; 135544 MW; D10A9D04115F532D CRC64;

Query Match 8.3%; Score 572; DB 1; Length 1220;
Best Local Similarity 22.4%; Pred. No. 3.4e-31;
Matches 270; Conservative 174; Mismatches 431; Indels 328; Gaps 43;

QY 321 KGTSLSKLSTHTLLGQPFQGWTVASWPLTILVLVSVVVALAAGLVFTLTTPV 380
Db 51 KGAVGQKAPLWIRARFOALFLSLGCHIQRHCGKVLFGILVFGALSVGLRAALETDS 110
QY 381 ELWSAPNSQARSEKAFHDQHPG-PFFRTNQVILTAPNRSSVRYDSLLLGPKNFGILDLD 439
Db 111 KLVWEGSRVSKELRYTYKEQGEESVFTSQMLIQTPKQEG-----TNILTQE 157
QY 440 LLLLELLEQLERHLQVNSPEAQNISLQDTCYAPLNP--DNT-----SLYDCCINSL 490
Db 158 AL--LLHLEALSASKVQSLYLGKSWDLNKICFKSGVPIIENVMTERIDKLFPCMVTP 215
QY 491 LQYFQNRTL-----LLLTANO--TLMQTSQVQDKHFLYCA 526
Db 216 LDCFWEGSKGGGSAYLFGMPDIIQMNLDPLUKMEELSQFTSLGFRMLDKAQVGHAYM 275
QY 527 NAPLTFKDGDTALALSCMADYCAP-----VFPFLAI---GGYKG----- 561
Db 276 NRPCLDPSDT---DC--PHSAPNKPQVQVFNIAELQGGCHGSKKFMHQEELLGER 329
QY 562 -KD-----YSEARALIMTFSL-----NNYPAGDPRLAQAKL-----WEEAFLEE 599
Db 330 VKDSQNALQSAEALQTFWLLSPKQLYEHFKDDYEIHDINWNEKATAILSWORKPVEV 389
QY 600 MRAP--QRRMAGMFQVTFTAERSLEDEINRTAEDLPFATSYIVIFLYISIALGSSSW 657
Db 390 VHGISPQNSSNNVYAFST---TLNDIMKSFDSVSVIRVAGVYLLMLAYAVTM---LRW 443
QY 658 SRVMVDSKATLGLGGVAVVLGAVNMAAGFFSYLGIRSSLVILQVVPVFLVLSVGADNIFIF 717
```


Db 659 ---PVQLRHVGSANQSRAMRLFRKLAEE--TSLEQ-----TGVYADFFQ-----FAEQYNA 703

QY 1108 ILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLLSIVMLVDTVGFWALWDSINAVS 1167

Db 704 VLPGLTSSIAVAGVAVVAVSILIPPEVAS-LWVSFSIVSNIGILGFTMTFWSVRLDFIS 762

QY 1168 LINLVSAGVMSVEFVSTRTSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVL 1227

Db 763 MVTIWSVIGFCVDFAHLAYNPAKENDIDAPERMRNALYAVAGFILMSASST-IIGVSFM 821

QY 1228 GLAKAQLQIPIFFRLNLLITLGLLHGLVFLPVLIS--YVGDVNPALALQKRAEAVA 1285

Db 822 ASAESVYFR-SPLKTIILLVILLGALHGLVFLPVLISMFYCGG-----SSKKAKEHMD 872

QY 1286 AV--WVASCNPHPSRVST 1301

Db 873 AVDQKLOAQYNNPARTAS 890

RESULT 83

Q9V9L4 DROME

ID Q9V9L4 DROME PRELIMINARY; PRT; 1169 AA.

AC Q9V9L4; Q86P36;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE CG11212-PA (R45036p)

GN Name=Pr; ORFNames=CG11212;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle K., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Foster C., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Mentrulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reineert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,

RA Svirskaas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195 (2000).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426065; PubMed=12537568;

RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,

RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,

RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

RT "Finishing a whole-genome shotgun: release 3 of the Drosophila

RT melanogaster euchromatic genome sequence."

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R.,

RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,

RA Ashburner M., Celniker S.E.;

RT "The transposable elements of the Drosophila melanogaster euchromatin:

RT a genomics perspective."

RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20 (2002).

RN [4]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a

RT systematic review."

RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).

RN [5]

RP NUCLEOTIDE SEQUENCE.

RG Berkeley Drosophila Genome Project;

RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,

RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskaas R., Smith E.,

RA Yu C., Rubin G.;

RT "Drosophila melanogaster release 4 sequence."

RL Submitted (MAR-2000) to the EMBL/GenBank/DBAJ databases.

RN [6]

RP NUCLEOTIDE SEQUENCE.

RG FlyBase;

RL Submitted (MAR-2005) to the EMBL/GenBank/DBAJ databases.

RN [7]

RP NUCLEOTIDE SEQUENCE.

RC STRAINS=Berkeley;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celniker S.;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBAJ databases.

CC -1- INTERACTION:

CC P09040:DBK; NbExp=1; IntAct=EBI-174538, EBI-1355576;

EMBL; AE003784; AAF57274.2; -; Genomic_DNA.

DR EMBL; BT003500; AAO39504.1; -; mRNA.

DR IntAct; Q9V9L4; -;

DR Ensembl; CG11212; Drosophila melanogaster.

DR FlyBase; FBgn0033068; CG11212.

DR FlyBase; FBgn0033068; ptr.

DR GO; GO:004872; F:receptor activity; NAS.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR006162; Ppantne S.

DR InterPro; IPR000731; SSD 5TM.

DR Pfam; PF02460; Patched; I.

DR PROSITE; PS00012; PHOSPHOPANTHINE; UNKNOWN_1.

DR PROSITE; PS0156; SSD; 1.

SS SEQUENCE 1169 AA; 133198 MW; F9D750BB015C1350 CRC64;

Query Match	8.1%	Score 562.5	DB 2	Length 1169
Best Local Similarity	21.5%	Pred. No. 1.5e-30		
Matches 219	Conservative 186	Mismatches 378	Indels 235	Gaps 35
Qy	322	GTSLSDKLSFSTHTLIGQFFQGMGTWASWPLTTLVLVSIVPVALAAGLVFETLTDPEV	381	
Db	4	GISCVDK-----TLNKSFYHLGICIAKHFGYFIIPVLLTLLCMGTGYOQLKYQIDPEY	56	
Qy	382	LWSAPNSQARSEKAFHDQHGFPPF--RTNQVILTAPNRSSRYRSDLLIGPKNFSGIL---	436	
Db	57	LFSPICEAGKTERAIVQYFKVYTHRENVGRIITPGR-----FGRVIVIT	102	
Qy	437	LDLLELLELQRLRLHQLWSPEAQRNISLQDICYAPLMPDNTSYDCCINSLLQYFQ	495	
Db	103	KGDENMIRREVFQELRQLD-----NI-IQN-ATTYDGDVTYKDCNARMENECFE	152	
Qy	496	NNRTLLLLTA-----NOTLMGQTSQYDWDKHFLYCANAPLTFKDGTLALSCWAD	545	
Db	153	ND--ILNLDALMDDIAGQLMLTFPFNFNPVTWDAHLF-----	188	
Qy	546	YGAPVFPFLAIGYKGKDYSEAEALIMTFSLNNYPAGDPRL-----AQAKL	591	
Db	189	--PVP-----PGGYK-----LTDENVISVPALQVYFVTDATKRODAGAE	228	
Qy	592	WEEAFLEEMRAFQRMAGMFO---VTFEASRLEDEINRTAEDELPIFATSIVY--IFLY	646	
Db	229	WEETFLRWVGAEN--SQPKHISVSYFASRTLQHELEKNTKTVVPYFSFTFLMLGLFSI	286	
Qy	647	ISLALGYSYSSWRVWVDSKATLGGVAVVLGAVMAAGFFSYIGIRSSLVILQVVPFLV	706	
Db	287	ITCMGMD-----AVRSKPPFLGMGNVSATMATLAAGLAMYCGI--EFIGINLAADFLM	338	
Qy	707	LSVGADNIFIPVLSEYQRLPRPGPREVHIGRALGRVAPSMLLCSLEATCFELGALTQM	766	
Db	339	IGIGIDTFVMLAGWRTKAK--MPVAERMGLMMSEAAVSTITSVDTFISFLIGIISPP	396	
Qy	767	PAVRTALTSLAVILDFLQMSAFALLSLDSKQBSASRLDVC--CCVQOELP-----	819	
Db	397	RSVRICTYSVFAVCFITLWHITFFAACMAISGYRERKNLHISIFGCRVQPMVAIKEBN	456	
Qy	820	-----PPQGGEGLLIGPPKAYAPFLHLHWITRGVWVLLFLALFGVSL	861	
Db	457	FLYKAIMAGGIDANDPNPINDKHMLMAFPKQXAAVINNKWKCAIILLAFASLYVGAC	516	
Qy	862	YSMCHISVGLDOELALPKDSVLLDYFLFLNRYE-----VGAPVYFVTTLYGNPSS	912	
Db	517	YGTQIQEGUERRKLREDSYVSVEFFDREDDYREPPYRMOVITAGPL-----NYS	568	
Qy	913	EAGMNAICSSAGCNFFGTQIQVATEFFEQSYLAIPASSWVDDFIDWLTPSSCCRLVIS	972	
Db	569	PLVQEQV-----ENLTSTLEHTSVTSRY-----TESWLSRFLSFLERN--	609	
Qy	973	GNPKDKFCPTVNSLNCNKCMISITMGSRVPSVEQFHXYLPWFILNDRPNIKCPKGGLAAY	1032	
Db	610	-----ELLNVTVDDQTFIDAVKEH--WLPPGNP-----F	637	
Qy	1033	STSVNLTSD--GOVLASRFMAYHKLNKSQDYTEALRAARELANITADLRKVPCTDP--A	1089	
Db	638	SLDVRFNEDETQIIATSFLLQAVNITDHEKENVR-----DLRQICKDSPLNA	686	
Qy	1090	FEVPPYITINVFYQYVITLPEGLFMLSCLVPTFFAVSVCLLGLDLSGLNLLSIVMIL	1149	
Db	687	SIFHPYP---VFPQFELVRPVSLOAMVIGAIMMIISFVIP--NILCSLWVAFSVISIE	742	
Qy	1150	VDTVGFMALMDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMG	1209	
Db	743	LGVAGYMALMDVNLDSISMINLINCIGFSVDFTAHICVYTMSSKKRSPKARVEALHSLG	802	
Qy	1210	SAVPAGVAMTNLPGLVIGLAKAQIQLIPFPRMLNLTITLGLHLGVFLPVLISYVGP	1267	
Db	803	LPIIQGSSSTIL-GIVALLAAQSYIF--LVFFKMWFLVFFGAMHGLFLLPVLISLFCP	858	

RESULT 84		Q4SP00_TETNG		Q4SP00_TETNG PRELIMINARY; PRT; 816 AA.	
ID	Q4SP00	Q4SP00	Q4SP00	Q4SP00	Q4SP00
AC	Q4SP00	Q4SP00	Q4SP00	Q4SP00	Q4SP00
DT	13-SEP-2005	(TrEMBLrel. 31, Created)	DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)	DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)
DE	Chromosome 15 SCAF4542, whole genome shotgun sequence.		DE	Chromosome 15 SCAF4542, whole genome shotgun sequence.	
GN	ORFNames=GSTENG00015058001;		GN	ORFNames=GSTENG00015058001;	
OS	Tetraodon nigroviridis (Green puffer).		OS	Tetraodon nigroviridis (Green puffer).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		OC	Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;		OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;	
OC	Tetraodontidae; Tetraodontidae; Tetraodon.		OC	Tetraodontidae; Tetraodontidae; Tetraodon.	
OX	NCBI_Taxid=99883;		OX	NCBI_Taxid=99883;	
RN	[1]		RN	[1]	
RP	NUCLEOTIDE SEQUENCE.		RP	NUCLEOTIDE SEQUENCE.	
RA	Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,		RA	Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,	
RA	Maucei E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,		RA	Maucei E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,	
RA	Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,		RA	Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,	
RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,		RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,	
RA	Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,		RA	Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,	
RA	Biemont C., Skalli Z., Cattolico L., Poullain J., De Bernardis V.,		RA	Biemont C., Skalli Z., Cattolico L., Poullain J., De Bernardis V.,	
RA	Cruaud C., Duprat S., Brotier P., Coutanceau J.P., Gouzy J.,		RA	Cruaud C., Duprat S., Brotier P., Coutanceau J.P., Gouzy J.,	
RA	Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,		RA	Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,	
RA	Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,		RA	Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,	
RA	Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,		RA	Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,	
RA	Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,		RA	Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,	
RA	Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,		RA	Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,	
RT	"Genome duplication in the teleost fish Tetraodon nigroviridis reveals		RT	"Genome duplication in the teleost fish Tetraodon nigroviridis reveals	
RT	the early vertebrate proto-karyotype.";		RT	the early vertebrate proto-karyotype.";	
RL	Nature 431:946-957(2004).		RL	Nature 431:946-957(2004).	
RN	[2]		RN	[2]	
RP	NUCLEOTIDE SEQUENCE.		RP	NUCLEOTIDE SEQUENCE.	
RG	Genoscope; Whitehead Institute Centre for Genome Research;		RG	Genoscope; Whitehead Institute Centre for Genome Research;	
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.		RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.	
CC	-1- CAUTION: The sequence shown here is derived from an		CC	-1- CAUTION: The sequence shown here is derived from an	
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is		CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is	
CC	preliminary data.		CC	preliminary data.	
DR	EMBL; CAAB0101542; CAF97632.1; -; Genomic DNA.		DR	EMBL; CAAB0101542; CAF97632.1; -; Genomic DNA.	
SQ	SEQUENCE 816 AA; 90662 MW; 2852CA728580E7D0 CRC64;		SQ	SEQUENCE 816 AA; 90662 MW; 2852CA728580E7D0 CRC64;	
Query Match 7.78; Score 532; DB 2; Length 816;					
Best Local Similarity 22.08; Pred. No. 1.4e-28;					
Matches 21; Conservative 160; Mismatches 385; Indels 202; Gaps 30;					
Qy	351	WPLTILVLSIPVVALAAGLVFTE--LTTDPVELWSAPNSQARSEKAFHQHGFPPFTN	408		
Db	13	WRRITAHLSGSDRSGLGFGFCFLQDRANDIEEQFTPDVQQAERYIQE---TFPGN	68		
Qy	409	QVILTAPNRSRYSDLSLLGPKNFGILDLDLLELLEQLERLHLQVSPQARNISLQ	468		
Db	69	ESMFRLRLSSGGNYATLIATSE--GDVLTFRVLQDIVELDSVRRVMV---HHDRSFEYQ	125		
Qy	469	DICVAPLNPDNTSLYDCINSLQYFQNNRTLLLTANTQTLMGQTSQVDKDHFLYCANA	528		
Db	126	DVACAGWG-----SCTPHNLDLVGDR-----ANLTF-----PW-----PHENG	161		
Qy	529	PLTFKQGTALALSCMADYGAPVPFFFLAIGGYKGQYSEAEALIMTFSLNNYPAGDPRLAQ	588		
Db	162	SVPLHGG-----LGGVKLWANSVVSQSCRAIRLFYTLRQDDR-TK	200		
Qy	589	AKLWEERAFLEEMAFQRRMAGMQVYTFARSLEDEINRTAEDLPFATSYIVIFLYIS	648		
Db	201	TDLWLOSFL---RLVSNASSASIRVSYSRSIQWEPKQTPGSLVCLFSAAYA-----IA	252		
Qy	649	LALGSYSWSRVMVDSKATIGLGGVAVLGAVMAMGFFSYGLRSLSLVLQVPPFLVLS	708		
Db	253	ITTSIVTCRWDSVRTKVVVALGGVSTALVLSGGFALLLLG-RPFWMTAASCP			

```
Db 312 VGLDDMFILISCRWRTRVLDSPRR-----LADTYGEAGVSISITTLTNALALLVGY 363
Qy 763 LTPPAVTEFALTSLAVILDFLLQMSAFVALLSLDSKRQEARSLDVCCCKVPQELPPG 822
Db 364 SSPFGSVNSFCYAGSVFCYLCVTFPLGACMALNGRRQAODHWFTC-----RKVPEDS 419
Qy 823 QGE-----GLLQG-----FFOKAYAPFLLHWTTRGVVLLLFALFGV 859
Db 420 AERTNISRICLGGRRGDITEMETEAMTDIFEKFGYGFLLTHSKVAKVLLVAGYLAA 479
Qy 860 SLYSMCHTSVGLDQELALPKDSYLDYFLFNRYFEVGAPYFVFTTLGYNFSSBAGMNAI 919
Db 480 STYGCLILKEGLEIKNLVLDSDYIIPYLEDQKH-----GEYGFNVN 522
Qy 920 CSSAGCNFSFTOKIYATEPEQSYLAIPASSWVDDFDMLTPSSCCRLYISGNKDKF 979
Db 523 -----VAKQPLLYWDQSEQ-----RLH-----S 542
Qy 980 CPSTVNSLNCNKNCSITMGSVR-----PSVEQPHKYLPMFLNDRPNIKCPKGGLA 1030
Db 543 CVSRFEGLTFFVSGTSLSWFLSQRYSNTNRLDVTSGAERTHLTHLEANP----- 592
Qy 1031 AYSTSVNLTSGQVLASRFMAHYKPLKNSQDYTEALR-AARELANITADLRKVPGTDP 1089
Db 593 VFKQDISFSAGQIQASRFFIQTQKIPKEEVMVTLRQTABECPLQLLV-----FHPA 645
Qy 1090 FEVP-PYITITNVFYEQLTILPEGLFMLSCLIVPTFAVSCILLGLDLSGLNLSIWM 1148
Db 646 FYLDQYTVVTAKTQVAVLVAVALVLSLALIPSLCSAW-----AFVCSV 694
Qy 1149 LVDTVGFMALWDISYNAVSLINLSAVGMSVEFVSHIRSPAISTKPTWLBRAKEATISM 1208
Db 695 MGVNAGFVALMGVNLDSVMSNLWCTGFSVDFSAHSVAFVSSSKTDIDGKAVELARL 754
Qy 1209 GSAVFAGVAMTNLPGLVGLAKAQLIQIFFRMLNLTLLGLHGLVFLPVLISYVG 1266
Db 755 GYVPLQG-ALSTILGWLLSGSVIFRTFKIIFLVIT-FGLIHSIVFIPVFLTLIG 810

RESULT 85
Q9XYP5_JUNCO PRELIMINARY; PRT; 1318 AA.
AC Q9XYP5;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative hedgehog receptor.
GN Name-ptc;
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrypsia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX NCBI_TaxId=39708;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99115913; PubMed=9915699; DOI=10.1126/science.283.5401.532;
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RA "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
  evolution."
RL Science 283:532-534(1999).
DR EMBL; AF117898; AAD31595.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
KW Receptor.
```

```
SQ SEQUENCE 1318 AA; 146474 MW; BA774C7133A22221 CRC64;
Query Match 7.6%; Score 522.5; DB 2; Length 1318;
Best Local Similarity 22.3%; Pred. No. 1.2e-27;
Matches 259; Conservative 173; Mismatches 421; Indels 307; Gaps 47;
Qy 336 LLGFOFGWGTWASWPELTILVLSVIPVVALAAGLVFTELTDPVELMASPNSQARSKA 395
Db 72 ILGCFLOGDAGKVL--VAIIVLSTFCV-----GLKSAQIHTRDVQLMVGGRLAEALK 124
Qy 396 FHDHQFPP-PRTNQVILTPNRRSYRDSLLLOPKNPFSGILDLDDLLLELLELQERLHL 454
Db 125 YTAQALGEADSSHTQLVI---QTKADPDVSLHP---GAL-----LEHL 162
Qy 455 QVMSPEAQRNI-----SLQDICYAPLNPD-----NTSLYD-----CCINSLLYQFQNN 497
Db 163 KVHAATRTVTHMYDIEWRLKDLCSIPDPFEGYHHIESIIDNVIPCAITPDLDCFWE 222
Qy 498 RTLL-----LLTANQTLMGQTSQ--- 515
Db 223 SKLIGDPYPIVPHLKHKLQWTHLNPLEVVEVKLKFQPPPLSTIEAVMKRAGITSAYMK 282
Qy 516 ---VDWKDFLYC-ANAPLTFKQGTALALSCHADYGAVPFF-----LAIGYKGD 563
Db 283 KPCLDPTDP--HCPATAP-NKKSCHI PDVAELSHGCGYFAAAYMHMPEQILVGGATRNS 339
Qy 564 YS---EAEALIMTFSL-----NNYPAGDPRLAOKLWEBAFLBEMRAFORRMAGMFOVT 614
Db 340 TSAURSARALQTVVQLMEREMYWADYKHQI GWNQEKAAAVLDQWQKFAEVRKI 399
Qy 615 FT-----ABRSLEDEINRTTAEDLPFATSYIVIFLYISIALGSYSWSRV 661
Db 400 TTSGSVSSAYSFYFSTSTLNDILGKFEVSLKNIILGYMFMILYVAVTL---IQW-RDP 455
Qy 662 VDSKATILGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVFPFLVLSVGADNIFIVLEY 721
Db 456 IRSQGVGIAGVLLSITVAAGLGCALLGFPFNASTQIVPFLALGLGVQDMFLTHTY 515
Qy 722 QRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTALTSLGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGLSVLLASLCNVMAFLAALLFIPAFVFCLOAAILLL 573
Qy 782 LDFLQMSAFVALLSLDSKRQEARSLDVCCCKVPQELPP----- 820
Db 574 FNLGSILLVFPAMISLDRRESAARADLLCCLMPESPPLPKKI PERAKTRKNDKTHRIDT 633
Qy 821 -----PCQGEGL-----LLGFFQKAYAPFLLHWTTRGVVLLLFALFGVLSYMC 865
Db 634 TRQPLDPDVSENVTKTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTSVWGAT 693
Qy 866 HISVGLDQELALPKDSYLLDYFLFN---YFEVGAPYFVFTTLGYNFSSBAGMNAICSS 922
Db 694 KVDGLDLDIVPENT---DEHFLSRQEKYFGF-----YN-----MYAVTQ- 732
Qy 923 AGCNFSP--TQKIYATEPEQSYLAIP-----ASSWVDDFDLWL----- 961
Db 733 ---GNFEYPTNQKLLY--EYHDQ-FVRIPNIIKNDNGGLTKFWLSLFRDLWLDLQVAFDK 786
Qy 962 -TPSSCCRLYISGNKDKFCPSTVNSLNCNKNCSITMGSVRPSVEQ----- 1007
Db 787 EVASGCI-----TQEWCKNASDE-GILAYKLMVQTGHVDNPFIDKSLITAGHRLVDK 838
Qy 1008 -----PHKYL-PWFLNDRPNIKCPKGL-----AAYSTSVNLTSGQVLASRPM 1050
Db 839 GIINPKAFYNLSAWATNDALAYGASQNLKAPQQRWHSPEDEVHLEIKSSPLIYQLP 898
Qy 1051 AYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAFEVPPYITITNVFYEQLTILP 1110
Db 899 FYLSGLSDTDSIKTILRSVRDLCLKYEA--KGLPN-----FPSGIPFLFWEQYL-YLR 948
Qy 1111 EGLFMLSLCLVPTFAVSCILLGLDLSGLNLSIWMILVDTVGFPMALWDSYNAVSLIN 1170
Db 949 TSLILALACALGAVPIAVVLLNNAWAAVLTALATLVLQLLGVALLGVKL SAMPVVL 1008
```


RX MEDLINE=90058658; PubMed=2582494; DOI=10.1016/0092-8674(89)90021-4;
RA Hooper J.E., Scott M.P.;
RT "The Drosophila patched gene encodes a putative membrane protein
RT required for segmental patterning.";
RL Cell 59:751-765(1989).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90015164; PubMed=2797178; DOI=10.1038/341508a0;
RA Nakano Y., Guerrero I., Hidalgo A., Taylor A., Whittle J.R.S.,
RA Ingham P.W.;
RT "A protein with several possible membrane-spanning domains encoded by
RT the Drosophila segment polarity gene patched.";
RL Nature 341:508-513(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blasej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall A.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP FUNCTION: Segmentation polarity protein. Acts as a receptor for
CC the hedgehog protein (HH). Associates with the smoothened protein
CC (SMO) to transduce the hedgehog signal leading to the activation
CC of wingless, decapentaplegic and patched itself. Participates in
CC cell interactions that establish pattern within the segment and

CC the imaginal disks during development. In the absence of HH,
CC represses the constitutive signaling activity of smo through fused
CC (FU).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the patched family.
CC -1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M28999; AAA28696.1; -; Genomic DNA.
DR EMBL; M28418; AAA28696.1; JOINED; Genomic DNA.
DR EMBL; X17558; CAA35591.1; -; mRNA.
DR EMBL; AE003836; AAF59062.1; -; Genomic DNA.
DR PIR; S06119; S06119.
DR FlyBase; FBgn0003892; ptc.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0008158; P:hedgehog receptor activity; IMP.
DR GO; GO:0004888; P:transmembrane receptor activity; NAS.
DR GO; GO:0004899; P:anterior/posterior lineage restriction, ima. . .; TAS.
DR GO; GO:0001746; P:Bohlig's organ morphogenesis; IMP.
DR GO; GO:0009880; P:embryonic pattern specification; TAS.
DR GO; GO:0007455; P:eye-antennal disc metamorphosis; IMP.
DR GO; GO:0042078; P:garm-line stem cell division; TAS.
DR GO; GO:0045877; P:negative regulation of smoothened activity; TAS.
DR GO; GO:0030707; P:ovarian follicle cell development (sensu in. . .; TAS.
DR GO; GO:0007432; P:peripheral nervous system development; TAS.
DR GO; GO:0007346; P:regulation of mitotic cell cycle; IMP.
DR GO; GO:0042306; P:regulation of protein-nucleus import; IMP.
DR GO; GO:0007224; P:smoothened signaling pathway; IMP.
DR GO; GO:0048103; P:smoothed stem cell division; TAS.
DR GO; GO:0048103; P:wing disc anterior/posterior pattern formation; TAS.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR Pfam; PF02460; Patched; I.
DR TIGRfam; TIGR00918; 2A060602; 1.
DR PROSITE; PS0156; SSD; 1.
KW Developmental protein; Glycoprotein; Receptor;
KW Segmentation polarity protein; Transmembrane;
FT TOPO_DOM 1 76
FT TRANSMEM 77 92
FT TOPO_DOM 93 427
FT TRANSMEM 428 448
FT TOPO_DOM 449 465
FT TRANSMEM 466 486
FT TOPO_DOM 487 492
FT TRANSMEM 493 511
FT TOPO_DOM 512 532
FT TRANSMEM 533 553
FT TOPO_DOM 554 562
FT TRANSMEM 563 583
FT TOPO_DOM 584 677
FT TRANSMEM 678 699
FT TOPO_DOM 700 931
FT TRANSMEM 932 952
FT TOPO_DOM 953 955
FT TRANSMEM 956 976
FT TOPO_DOM 977 1007
FT TRANSMEM 1008 1028
FT TOPO_DOM 1029 1056
FT TRANSMEM 1057 1077
FT TOPO_DOM 1078 1082
FT TRANSMEM 1083 1103
FT TOPO_DOM 1104 1286
FT DOMAIN 1287 1332
FT COMEBIAS 1227 1332
FT CARBOHYD 142 142
FT CARBOHYD 298 298

FT CARBOHYD 335 335 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 388 388 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 807 807 N-linked (GlcNAc...) (Potential).
FT CONFLICT 111 111 R -> G (in Ref. 2).
FT CONFLICT 274 274 G -> A (in Ref. 2).
FT CONFLICT 332 332 A -> R (in Ref. 1).
FT CONFLICT 636 636 P -> A (in Ref. 2).
FT CONFLICT 862 866 DVF -> ASSPTELLRANCIRNR (in Ref. 2).
FT CONFLICT 866 866 Y -> N (in Ref. 2).
SQ SEQUENCE 1286 AA; 5P22A956F8BE0EC6 CRC64;

Query Match 6.9%; Score 474; DB 1; Length 1286;
Best Local Similarity 20.4%; Pred. No. 3.2e-24;
Matches 245; Conservative 177; Mismatches 457; Indels 322; Gaps 42;

QY 321 KGTSLDLSFSTHTLGGFGWGTWASWELTILVLSVIPVVALAAGLVTELTTPDV 380
DB 43 KKGARGRTAIYLRVFSQSHLETSSVQKHAGLVFVAILVLSVFCVGLKSAQIHSKVH 102
QY 381 ELWAPNSQARSEKAFHQHFG-PFRFTNQVILTA---PNRSSYRYSLLGPKNFGIL 436
DB 103 QLWIOGGRLEALAYTKTIGEDSATHQLLIQTTHDENASVLPQALL-----A 153
QY 437 DLDLLELLELLELQERLHLQVMSPEAQRNLSLODICVAPLNPNTSLY-----DCC 486
DB 154 HLEVLVKATAVVKHLYDTE-W-----GLRDMCMSPSPFEGIYIEQILRHLIPCS 204
QY 487 INSLLOYFON-----NRTLLLLLTAN-----QTLMGQTSQVDWKDH 521
DB 205 IITPDLDFWEGQLGPESAVVPGNLQRLMTTLNPASVMYQMKQSEKISDFETV 264
QY 522 FLYCANAPLTFKGTGA-LALSQW-----ADYCAPVFPPLAIG--GY 559
DB 265 EQMKRAAI-----GSGYMEKPCLPNPNCPDPAKNKSTQPPDGA---ILSGCYGY 316
QY 560 KGOYSEBALIMTSLNPNYPAGDRLAQA-----589
DB 317 AAKHMHPEELVGGAKRN-RSGHLRKAQALQSVVQLMTEKEMYDQWQNYKVHHLGWTQ 375
QY 590 -----KLWEEAF---LEEMAFORMMAGMQVFTFAERSLEDEINRTTAEDLPIFATS 639
DB 376 EKAAEVLANQRNFREVEQLRKQSRATNTDYIVFSSAALDDILAKFSPHSALSIVIG 435
QY 640 YIVIFLYTSLALGSSYSSSRVMDSKATIGLGGVAVVLGAVMAAMGFYSILGIRSLVIL 699
DB 436 VAVTVLY---AFTLLRW-RDPVQGSSVGAVGLLMCFSTAGLSALLGIVFNAAST 491
QY 700 QVPELVLSVGNADNIFIVLEYORLPRRGPPEVHIGRALGRVAPSMMLCSLSBAICFF 759
DB 492 QVPELALGLGVDHIFMLTAAYAESNR--EQTKL-----ILKKVGFSILFSACSTAGSFF 545
QY 760 LGALTPMPAVRFTALTSGTGLAVILDFLLQMSAFVALLSLDSKQEAASRLDV-CCCV----- 813
DB 546 AAATFVPVPAKVFCLQAAIVMCSNLAAALLVPAMISLDLRRRTAGRADIFCCCFPVWKE 605
QY 814 KQELPPL-----PGQEGEL-----LLGF 831
DB 606 QPKVAPPVLPNNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGSHSLASLSLATP 665
QY 832 FOKAYAPFLHWTIRGVVLLFLALFGVLSYGMCHI SVGLDQELAPKDSYLLDYFLFLN 891
DB 666 AFQHTVTPFLMRSWKFLVWGLFALALISLYASTRLQDGLDIIDLVPKDSNEHKEFLDAQ 725
QY 892 RYFEGAPVYFTTLYGYNPESSEAGNAICSSAGCNFFSTQKIQVATFPEQSYLAIP-- 949
DB 726 RLFL--GFYSMAVATQ-----NFEYPTQQQLRDY-HDSFVRVPHV 763
QY 950 -----ASSVDDFDLWLTSPSCRLVYISGPNKQKFCFSTVNSLNCNKC----- 993
DB 764 IKVNGGLPDPFWLLLFSEWL-----GNLQKIFDEYRDGRLTKECWFNASSD 811
QY 994 -----MSITMGSVRPSVEQ-----FHKYL--PWFILNDRPNKCPKG 1027

DB 812 AILAVKLIVQTHGVNDPVDKELVLNRLVNSDGIINQRAFYNYLSAMATND----- 862
QY 1028 GLAAYSTSVNTSDQVLAASRSMAYHKPKNSQD-----YTEALRAARE 1071
DB 863 -VFAVG-----ASQGLYPEPRQYFHP--NEYDLKPKSLPVYAQMFPYLHGLTDSQ 914
QY 1072 LAANTADLRKVPGTDPAFEV--FPYTTNVFYEQYVLTILPEGLPMLSLCLVPTFAVSL 1129
DB 915 1-KTLIGHIRDLVSVKYGFGFLPNYPSGPIPFWEQYMT-LRSSLAMILACVLLAALVLS 972
QY 1130 LGLDLRSGLNLNLSIVMLVDTVGFMAWLDISYNAVSLINLVSAVMSVPEVSHITRSP 1189
DB 973 LLLLSVAAVILSVLASLAQIFGAMTLLGKLSAIPAVILSVGMMLCP--NVLLSL 1030
QY 1130 AISTPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQITQIFFFRNLILITLL 1249
DB 1031 GPMTSVGNRRQVQLSMQSLGPLVHGMLTSGVAVFMLSTSPFVIRHFVLLVLCV 1090
QY 1250 GLLHGLVLPVLVLSVGPDPNPALALOKRAEEAAVAVMVASCPNHPSPRSVADNIYVNH 1309
DB 1091 GACNSLLVFPILLSMVGPEAB-LVPLE-----HPDRISTPSPLPVRS 1131
QY 1310 S 1310
DB 1132 S 1132

RESULT 90
Q09938 CABEL
ID Q09938 CABEL PRELIMINARY; PRT; 1015 AA.
AC Q09938;
DC 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein C54A12.1.
GN Name=ptr-6; ORFNames=C54A12.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; U28733; AAA68302.1; -; Genomic_DNA.
DR PIR; T15830; T15830.
DR Ensembl; C54A12.1; Caenorhabditis elegans.
DR WormBase; WBGene0004221; ptr-6.
DR WormPep; C54A12.1; CE01865.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003392; F:hedghehog receptor activity; IEA.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; I.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1015 AA; 116343 MW; 6D69AF443028F6C5 CRC64;

Query Match 6.9%; Score 473.5; DB 2; Length 1015;
Best Local Similarity 19.2%; Pred. No. 2.6e-24;
Matches 226; Conservative 186; Mismatches 417; Indels 347; Gaps 38;

QY 289 IILCSVFAV---VTILLVGFVAPADSKMVKPKGTSLSKLSFSTHTLGGFGW 345
DB 11 LLKTSIFENGKVVLLALIAHRF-PNRCRIPYLDKK---LEKVFANYTKHLLVDY-- 61
QY 346 TWVASWPTILVLSVIPVVALAAGLVF-TELT-----DVV-----ELWSA 385
DB 62 -----YVFLIAPVILTACSGFYWKELTLLDARKLYTPVSAPSKEQVFSLEMPV 114

```
QY 386 PMSQARSEKAFDQHGFGFPRTNQVILTPNRSYRYDSLLGPKNFGSGLDLDLLELL 445
Db 115 KSYEFLPERT-----FQWRYLYLVHGRPYENG-----YFNILEGSLYNDIA 158
QY 446 ELQERL-----RHLQWSPQAQRNLSLODICYAPLNPNONTSLYD 484
Db 159 KIEBELVTNVFPREAALEDNAPLMRKHLNTEIEENRITI-----TPKD 205
QY 485 CCINSLLOYFQNNRTLLLTANQTLMGQTSQVDWKDFLYCANAPLTPFKDGTALALSCMA 544
Db 206 VCMWYGDYREKNVIELKRRHNL-----KRGISVTFPOVN 243
QY 545 DYGAPEVFPPLAIG--YKGRDYSABEALIMTFSLNYPAGDPRLAQAKLWE--EAPLE 598
Db 244 QEGPTIYAFVIGGVDTPNDTIKARAWRLWFLKFDDEQEQL--AKFWEDTAERFVR 301
QY 599 EMRAPQRMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYVIFLYISIALGSYS--- 655
Db 302 ETYADHPTI-----QCHIKHSRIVDQGLTENANRLKPYENVTVIAVLVFTAF-----YSVKW 353
QY 656 -----SWSRWVD---SKATLIGGGVAVILGAVNAAMGFSYLGIRSLVILQVVPVLV 706
Db 354 YFRMDHWP-LHIDWLRKPMALGGLSVLAILSGIGLLWFGMFPFAEITL-IAPPLV 411
QY 707 LSVGADNIFIFVLEY-----QRLPREPGPREVHIGRALGRVAPSMLLCSLSEAICFFLGA 762
Db 412 LSGVDDMFIAVAWHNTEMYPGSPVKMKRMIEAMSESAVALIFTSFDVLSFGAGT 471
QY 763 LTPMPAVTFTALTSGLAVILDFLLQMSAFVALLSDSKRQASRLDVCVV----- 813
Db 472 ITDIIAVOGFCAMTAACTFFFLYQITFFAALMWISAKAQSGRNSCMPCITAGDIYIE 531
QY 814 -----KQELUPP 820
Db 532 DCSLPNLKKKSKSRDKAKAEKKEKBAKNDKNMEIEBENAENAEKSSYSDSPSPQIHI 591
QY 821 PCQSGLLGFGOKAYAPFLHWHITRGVVLILLFALFGVLSYSMCHISVGLDDELALPKD 880
Db 592 PKVSRGAMGHFPRODYVFWLWMLNWKLPFMPVTFIYLGISVVICVMEQGLDYDKLLHS 651
QY 881 SYLLDYFLFNRYFVGAPVYVFTVTLGYNFSSEAGMNAICSSAGCNFSTQKIQYATEF 940
Db 652 DPLVE--ALKKEIEL-----FHGGDQIEIA--- 674
QY 941 PEQSYLAIPASSWVDDFIDWLTTPSSCCRLYISGPNKDFCFSTVNSLCLNKMCSITMGS 1000
Db 675 -----IQNCPNITIAE 685
QY 1001 VRPSVEQFHXYLPWFLNDRPNIKCPKGG-----LAAYSTSVNLTSD-----GQ 1043
Db 686 SDRIE-----LVAQEFENISYSLGKGSGTFSWMREYKYSNLTGSGYLNDRNSWIVGV 738
QY 1044 VLASRFMAYHK-----PLKNSQDY-TEALRAAR-----ELAAINITADLRKVP 1084
Db 739 YWMSQLFAYKLSQDFWANESDYDTLELKSRYFRIGVHRLSTPTDLVL-ITEBLRGA 797
QY 1085 GTDPAPFVPPYITNVFVEQYITLPEGLFMLSCLCLVPTFAVSCLLGLDLRSLNLNLS 1144
Db 798 DRHPDLNIVTYQOSRAIADQLNVLSTTINDTLAMFCMFCVALLFIENPI-CALFITFA 856
QY 1145 IVMILVDVTGFMALDISVNAVSLINYSAGMSVEFVSHSTRSPAIS-TRKPTWLERAKE 1203
Db 857 MYTIDIGVIGFSLSVKLDPSIMITIIIMSIGFSIEFSAHIITHGVFVNSDNLSPAEDRCD 916
QY 1204 ATISMGSAVFAVAMTNLPGIIVLGLAKAQLIQIEFFRNLIKITLLGLHGVFLVPLVLS 1263
Db 917 AMEKLAWPVHG-SLSTILGVFLAFIDSVMVLFVFKTISLVL-IIGAWHALMLLPIILLS 974
QY 1264 YVGP-----DVNPALALEQKRAEBAVAMVASCP 1293
Db 975 MCPIVIERLSDASKASDRRRRLKENQNSVYAINLP 1010
```

```
RESULT 91
Q61PCI CAEBR
ID Q61PCI_CAEBR PRELIMINARY; PRT; 915 AA.
AC Q61PCI;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG07607.
GN Name=CBG07607;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000032; CAB63242.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedghehog receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Hypothetical protein.
SQ
SEQUENCE 915 AA; 103951 MW; DD7AB4E5B0125E50 CRC64;

Query Match 6.8%; Score 470.5; DB 2; Length 915;
Best Local Similarity 20.2%; Pred. No. 3.6e-24;
Matches 208; Conservative 195; Mismatches 391; Indels 235; Gaps 32;

QY 336 LLGQFQCGWGWASWPLTILVLSVIVVLAAGLVFTE--LTDPVVLWGNAPNSQARSE 393
Db 9 LVNRAFVYGPVHWRWFCFISFLFTMACSIGFLRWKDLRVDDPAYVTFPSDARWRE 68
QY 394 KAFDQHF-----GPFRTNQVILTAPNRSYRYDSLLGPKNFGSGLDLDLLELL 444
Db 69 INVFNENWPLDENKFLPGKSPETK-----RVNIIIRAKDGGSIMRDNVLOBI 116
QY 445 LELQE-RLRHLQWSPQAQRNLSLODICYAPLNPNONTSLYD--CCINSLLOYFQNNRTLL 501
Db 117 QILANNWIMNNTSVPTDLDLKFNLTYQDLCLS-----YDWVCGANBHIQMLIRNDV- 166
QY 502 LLTANQTLMGQTSQVDWKDFLYCANAPLTPFKDGTALALSCMADYG-APVFPFLAIGGYK 560
Db 167 ----NQVLDLHFPFGTKDTVPYLSG----IFGDVNFENGTLSDAKLTQLFVFLKQDQNM 219
QY 561 GKDYSEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEMRAFORRMAGMFOVTTAERS 620
Db 220 VVEYSKFSYAIEFLNQVYSSD-----VISLSFAHYQS 253
QY 621 LEDEINRTTADLPIFATSYVIFLYISIALGSYSWSRWVD----SKATLGLGVAVVL 677
Db 254 LEDGLDENAKAFVNFVVSFFFLAMAYALISSFLSKSSNAKKIDWTISSRPLWATAGMFTTV 313
QY 678 GAVMAAGFFSYGIRSLVILQVVPFLVSGADNIFVLEVORLPRRPGCEPREVHIG 737
Db 314 LSIVSAPGFLGLGVHYN-VINTIIPFLIIAIGDDMFLMNCWQDTSKLSVPE--RMS 370
QY 738 RALGRVAPSMLLCSLSEAICFFLGTALTPMPAVRTFALTSGLAVILDFLLQMSAFVALLSL 797
Db 371 KTLSHAGVAVITTNVTDIMSFAIGCITDLPGLQPFICVACVSVAPSYLYQLTFFSGAMAI 430
QY 798 -----DSKRQE----- 803
Db 431 MGEVEREKRHCLFFYRTLRLVDISKNNEADSKLQOIKNRSASPAPFNLYSSNSFSFSSD 490
```

QY 804 -----ASRLDVCCVKQELPPPG-----QGEGLLGFQKAYAPFLHWT 845
 Db 491 TDSFSKKTIPAEFAWKQSPNSISEKSKSHGPHBEKDRIVYFLGQVYGPFLSNSV 550
 QY 846 RGVLILLFLALFVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYEFGVAVFT- 904
 Db 551 RLPSAVIFALYIALYALYGCNREGNPGNLTNDHYIAKYSDIXHFWRIQPOLHVAVL 610
 QY 905 -----TLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFD 959
 Db 611 NPKLTLGEN--REKLLDVVKS---PENTPYTLG-REGTVF-----FLLEYVN 652
 QY 960 WLTPSSCCCLYISGPNKDFCSTVNSLNCNCSITMGSVRPSVEQFHKYLPHFLNDR 1019
 Db 653 YLTD-----LNVEIADTEKLMNTKLS-----WLKYTGSGTQ-----WASNIR 690
 QY 1020 PNKCPKGLAAYSTSVNLTSQOVLASRFMAHYHKLKNSQDVTEALRAARELANITAD 1079
 Db 691 MN-----GTEFOSFRFQALKNFVFNPDHKKHAKLLREIAD----- 726
 QY 1080 LRKVGTDPAFVFPVYTIINVFYQYLITLP---EGLFMLSCLVPTFAVSCLLGLDLR 1136
 Db 727 -RQPNVNVVYHEAFP-----FADQYLIILPATIQNVVISLLCM---AVVSFLVPSLP 775
 QY 1137 SGLNLLSIVMLVDVTPMALWDISYNAVSLINLSVAGMSVEFVSHITRFAISTKET 1196
 Db 776 SGFVFSVSVISNIGVFGYMTLWGNLDVMSISIMSGFADLSAHIYAFVTSHGDT 835
 QY 1197 WLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQLQIIFERLNLITLGLLHGLV 1256
 Db 836 -KQVIGALETLQWPIFOGASST-IGAGISILYTVDAYII-LVFFKTIWLTMLGAIHGLF 892
 QY 1257 FLPVILSYV 1265
 Db 893 FIPVFLSVV 901

RESULT 92
 Q621T1 CAEBR
 ID Q621T1 CAEBR PRELIMINARY; PRT; 974 AA.
 AC Q621T1
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Hypothetical protein CBG02358.
 GN Name=CBG02358;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RG The C.briggsae Sequencing Consortium;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -! CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CAAC0100012; CAE59072.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD 5TM.
 DR Pfam; PF02460; Patched; 1.
 DR PROSITE; PS0156; SSD; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 974 AA; 112088 MW; 0286DD1D4A6A8892 CRC64;

Query Match 6.7%; Score 464; DB 2; Length 974;
 Best Local Similarity 19.7%; Pred. No. 1.1e-23;
 Matches 221; Conservative 173; Mismatches 386; Indels 344; Gaps 38;

QY 324 SLSDKLS-----FSTHTLGGQFQGMGTWASWPLTILVLSVIVVVALAAGLVF---TELT 376
 Db 7 TLDKKEFFAWYTRHLLVDYVY-----VFLIAPILTAICSGGFGYWKELT 53
 QY 377 T-----DPV-----ELWSPNSQARSEKAFHDQ-----HFGPFFFTNQV 410
 Db 54 LLDARKLYTPVSAPSWMREBEVSELPVKSYPFLPERTFQWNRYYLYLVHGRPFENGSY 113
 QY 411 IL-----TAPNRSYRYDSLLGPKNFSGILDLDLLELLELLELLELLELLEL 455
 Db 114 NLEGSYLLDIAKIBEEVATNVSFREALEDNPALM-----RGRHLN 156
 QY 456 WNSPRAQRNISODICYAPLNPNTSLYDCCNSLQYFQNNRTLLLLPANOTLMGQTSQ 515
 Db 157 STEIEVNQNTTI-----TFKDVCMNVGDCYREKNVIELKRRHELN----- 198
 QY 516 VDMKHFLYCANAPLTFKQGTALALSCMADYCAPVPPFLAIG---YKGDVSEAEALIM 572
 Db 199 -----KRGISVTFPQVQNEGTPIYAFVIGGVDTYPNDTIKYARAMEL 241
 QY 573 TFLNNYPAGDPRLAQAKLWE---BAFLERMAFORRMAGMFQVTFPERSLEDEINRTT 629
 Db 242 WYFLAKFDDDEBQOL--AKWEDTAKFYRE--TYEDHPT--IQCHKHSRIVDQGLTRNA 295
 QY 630 AEDLPFATSYIVIFLYISLALGYS-----SNRVMVDSKATLGLGVA 674
 Db 296 NRLKVFNVVTIAVLVLTAF-----YSVKWYFRTHCWPLHIDWLR-----SKPMALGGVL 347
 QY 675 VVLGAVMAAMGPFSLYGLIRSSLVILQVVPFLVLSVGADNIFIPVLEY-----QRLPRPGE 730
 Db 348 SSVTAILSGIGLLHFGMPFAETIL-IAPFLVLSIGVDDMFIAVAAMHNTMKYGPSPK 406
 QY 731 PREVHIGALGRVAPSMLLCSLSAICPFLGALTMPAVRTTALTSLGLAVILDFLLQMSA 790
 Db 407 VMKKRMVAMESAVAIPTSTFVLVSGVGIITDIIAVQGCAMTAACMFFFLYQITP 466
 QY 791 FVALSLDSKQSEASRLDVCCV-----KPOE----- 817
 Db 467 PAALMVISAQAMSGRNSCMPCTADDIYTIEDGSLQPNLKKSKSRKDAKEKPEKMN 526
 QY 818 -----LPPGQ-----GEGLLGFQKAYAPFLHWTGRGVLLFLAL 856
 Db 527 EIAESEIABKRSYSSDPFSQIHIPVKSRGAMGHFRDYVFWLLNWKTKLFLVFIYV 586
 QY 857 FGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYEFGVAVFTVTLGNFSEAGM 916
 Db 587 LGTAYIGCVMEQGLDYDKLLHSDPLVE---ALKREIDL----- 623
 QY 917 NAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFDLWLTSPSSCCCLYISGPNK 976
 Db 624 -----FHGGDQIEIA----- 633
 QY 977 DKPCSTVNSLNCNCSITMGSVRPSVEQFHKYLPHFLNDRPNKCPKGG-----IA 1030
 Db 634 -----IQCNPLTIPENRDIEM-----IAQEFENISYSLGKGTSFWMR 673
 QY 1031 AVYSTSVNLTS-----DQGVLA---SRFAYHKLKNSQDYTEALRA---ARELAN- 1075
 Db 674 EYKYSNLTSYLNDRNSWIVGVVWSQLFAFYKLM---SQDFVWANEADYDTLELKSFR 731
 QY 1076 -----ITADLRKVPETDPAFEVFPYTTINVFYQYLITLPEGLFMLSIC 1119
 Db 732 FRIGVHRLSTPTDLVLIITEELRGVADRPDLDDIVYQQSRAIDQLNVLSTIINDTLA 791
 QY 1120 LVPTFAVSCLLGLDLRSGLLNLSIVMLVDVTPMALWDISYNAVSLINLSVAGMSV 1179
 Db 792 MFCMFCVALIIPNPI-CALFITLANVTIDIGVIGLSLWSVKLDPISMITIIMSIGFSI 850
 QY 1180 EFVSHITRSPALS-TKPTWLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQLQIIF 1238
 Db 851 EFSAHITGFPVSNENNLAFRCVDMAMEKLAMPVVHVG--SLSTILGVFLAFIDSVMVLVF 909
 QY 1239 FFRNLNLTLLGLHGLVFLPVILSYVGVDPVNPALALEQKRAEE 1282

336 LAGQFQCGWGTWASWPLTILVLVSVIPVALAGLV-FTLE-TTDPVELWSAPNSQARSE 393
9 LVNAPFWGPVHWRWFCFISPLFLTLACSQVGFRRMTLRLVDDPSYFVTSDAWRRE 68
394 KAFDQHGFPFRTRNQVILTAAPNRSSRYDLSLLGPKNFSGILDLDLLELLEL-GERLR 452
69 ISVFENW-PL--DENKFLPGKSPKAKRFINILIRAKDGGSMRDVNLHEILNQWIMW 125
453 HLOWSPKQAINSLQDICYAPLNPONTSLYD--CCINSLLQYFQNNRTLLLTANTQTM 510
126 NISIPTDLLKFNLTQDCLLS-----YDWVCAGNEHIQWLLLRNDV-----NQILD 171
511 GOTSQVDWQHELYCANA---PLTPKDGTTALALSCWADYG-APVFPFLAIGYKGDYSE 566
172 LHPFRGGTKDTPVYLGIGFDGVQFQNGT-----LSDAKLTQLFYFLKQDKQWVEEYS 225
567 ABALIMTFLSNVYPAGDPRLAQAKLWEAPLEBRAPORMAGMFOVFTASRLEDEIN 626
226 K-----PSYAL-----EYFLNQVYS-----SDVITLSFAHYQSLEDGLD 259
627 RTAEDLPFATSYIVIFLYI---SLALGSYSWSRWVMDSKATLGLGAVVVLGAVMAA 683
260 ENAKAPVNPVNVFVFLAMVALVSSFTLKSSATKIDWISSKPLWLAAGMPSTVLIIISA 319
684 MGFPSYLGIRSSLVLQVVPFLV-----SVGADNIFIFVLEYQRLPRRP 728
320 FGFLFILGRVYN-VINTIIFLIGEFQSFQKMEKNQIIPAIGTDDMFLMNACWDQTSKL 378
729 GPREVHIGRALGRVAPSMLLCSLEAICFFGALTPMPAVRTEALTSLGLVILFLLQM 788
379 SVPE--RMSKTLSHAGVAIVTNTVDVMSFAIGCIITDLFGIOPFCIACVSAFSTFYQL 436
789 SAFVALLSL-----DSKRQE----- 803
437 TFFSGAMATGVEREKHCLFFRTFQLVDISKNERADSKLQIKSRSASPAPNYLS 496
804 -----ASRLDVC--CVKQELPPG-----QGEGLLGFQKAYAPPL 840
497 SSNSFSSDSDSFSSKKTIPABFAWKEQSPNSLSKSKSDEEKDRIVHFITKIGYPI 556
841 LHWITRGVVLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYEFGAPV 900
557 LNSVRIEFGLLFVVLAIAMTCTNFREGLPNGNLVNDHVIYAKFSDIKHFWIRGAQL 616
901 YFVTTLGVSFSEAGNNAICSSAGCNFSTOKIQVATEFPFQSYLAIPASSWVDVFDW 960
617 HVALNPPMLTISENNEL-----LKVSAPENTQY-TLREGTVFLEY 661
961 LTPSCCRILIYISGPNKDKPCSTVNSLNCNMSITMGSVRPSVEQFHYLPWFLNDRP 1020
662 LN-----YLSLNAE-----VEDTERLWTKLNSMLKY----- 689
1021 NIKCPGGLAAYSTSVNLT-SDQVLASRPMAYHKLKNSQDYTEALRAARELAN----- 1075
690 -----TGGSTQWASNLKINKTGDSPQAFRQIALKNFVEPNDKHAAQLLRDIADHPFN 744
1076 -----ITADLRKPGCTDPAFVFPVTINV-----FVEQYLTILP-----EGLFMLSCL 1120
745 VVVYHEVSFGNKKIL-NDFISHSYCVAQKNIPKLAFFPADQVLIILPATIQNVVLSLQM 803
1121 VETFAVSCLLGLDLSGLNLLSIVMLVDTVGFMAWDISYNVAVSLNLSVAGMSVE 1180
804 ----AVVSPLLVPSLPSGPIVFSIVSINIGVGYMTLWGNLDAVMSIIMSIGFAVD 859
1181 FVSHITRPAISTKTWLERAKEATISMGSAVAGVAMTNLPGLVILGLAKAQLQIPIFF 1240
860 LSAHIYAVFTVSHGDT-KQRVIGALETLGWPIFOGASST-IAGISILYTVDAIIL-LVFF 916
1241 RLNLTLITLGLLGLVLPVILS 1263
917 KTIWLTMLGATHGLFFIPIFUS 939

RESULT 95

Q6AW15 CABEL
ID Q6AW15 CABEL PRELIMINARY; PRT; 1358 AA.
AC Q6AW15;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Patched family protein 3, isoform a.
GN Name=ptc-3; ORFNames=Y110A2AL.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Eubaditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "genome sequence of the nematode C. elegans: a platform for investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AC024746; AAT92058.1; -; Genomic DNA.
DR Ensembl; Y110A2AL.8; Caenorhabditis elegans.
DR WormBase; WBGene0004210; ptc-3.
DR WormPep; Y110A2AL.8a; CE37163.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008159; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1358 AA; 152075 MW; 1E920618B75DF741 CRC64;
Query Match 6.6%; Score 455; DB 2; Length 1358;
Best Local Similarity 22.8%; Pred. No. 7.7e-23;
Matches 215; Conservative 146; Mismatches 373; Indels 209; Gaps 37;
QY 473 APLNPNTSLYDCCI---NSLIQYQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAP 529
DB 449 AKKNEKEAKOLLCLLEYGSSLLKWNQENPERL-----GEFLTKEEP 490
QY 530 LTFKDGDTALALSCWADYGAPVFPF---LAIGYK---GKDYSEARALIMTFLSNVYPAGD 583
DB 491 DYPNGYDVMTCGCKG-FGKKIMWEPDLIIGGIQDNGKLV-ARALQSVFLVSGAYDVF 548
QY 584 PRLAQAQKWEAPLEBRAPORMAG-----MPQVTFARSLSEDEINRTAEDLPFATS 639
DB 549 ARIKNDKTDSPGLDR-HHFQPMWAGEIISTWQNRFT-KRLYSHELNRERRQFHPLASTS 606
QY 640 -----YIVIFLYISLALGSYSWSRWVMDSKATLGLGAVVVLGAV 680
DB 607 IADMLEEFSQFNYIIIVIGYILMVYAAPTQGRFGW-WLAVQSNVALAICGVILTTIS 665
QY 681 MAAMGFFSYLGTIRSSLVLQVVPFLVSGADNIFIFVLEYQRLPRRPGEPREVIHIGRAL 740
DB 666 ICGLGATHLGINFNAATQVVPFLSLGLGIDDMELLNHYDEIINCNKN---EIGVLL 722
QY 741 GRVAPSMLLCSLEAICFFLGAITPMPAVRTPALTSLGLAVILDFLLQWSAFVALLSLDSK 800
DB 723 KETGMSVMLTISINNLAFISGYVLPALRSPSCQTALLAFNLFLFIPFAMIGIDLR 782
QY 801 RQEARSLDVCCVK-----POELP-----PPQGGEGLL 828
DB 793 KQKRRDLAYCSRGNPOMATSSQVPSNVSNMSRAELAGYEKQADEYKRHEPWTVTG-- 840
QY 829 LGFFOKAYAPFLHWTIRGVLLPLALFGVLSYSMCHISVGLDQELALPKDSYLLDLYPL 888
DB 841 -GFLNKIYIPALKNNVVKACVLITGTTAVFVGLYGMTSTLGLADVLPHPTPAFLR 899
QY 889 FLNRYFVFGAPVYFV-----TTLGYNFSEAGNNAICSSAGCNFSTOKIQVA 937
DB 900 AREQVFSF-YPMFAVLRGDKLDIPNQOQLIBEYRAQLG-----SSKFMK 943


```
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; U53149; AAA96120.1; -; Genomic_DNA.
DR PIR; A89153; A89153.
DR Ensembl; C24B5.3; Caenorhabditis elegans.
DR WormBase; WBGene0004216; ptr-1.
DR WormPep; C24B5.3; C805861.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR006162; Ppantne.S.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; I.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 956 AA; 109453 MW; 493A4C6BF4D26DCA CRC64;

Query Match 6.3%; Score 436; DB 2; Length 956;
Best Local Similarity 21.2%; Pred. No. 1.1e-21;
Matches 233; Conservative 177; Mismatches 353; Indels 338; Gaps 48;

QY 345 GTWASWPLTILVLSVPPVLAAG-LVFTLTT-DEVELMSAPNSQARSEKAFHDQ-- 400
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 400
QY 16 GENVHKFRLLYLLSPAILTFAFGLWFTETQTTNDPQVFPSPANAPWRYERAVLTERWP 75
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 75
QY 401 -----FGP-----FFRTNQVILTAPNRSSRYSDLSLLGPKNFSGILDLDLLELLEQLRLR 452
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 452
QY 76 LDEERFPWGSYDLHGVDVIASGRHPDY-----GRN-----ILNIRYLDEVARINDVII 127
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 127
QY 453 HLQVSPAEQRNLSL-----QDICYAPLNPNTSLYDCINSLLOQFQNNRTLLILT 504
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 504
QY 128 H-----NLTPVPDINGKHYDIAYT-----DLCMYDWACYLNDHTTMLMP 167
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 167
QY 505 ANQ-----TLMGQTSQVDKDHFLYCAN 527
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 527
QY 168 KTRWGNFSGPPAELASDIINTQVITYPICWRTGPIYFGALVGAPNLVDEGHFDYASA 227
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 227
QY 528 APLTP--KDGITALALSCMADYGAPVPPFLAIGYGK-----DYSEBALIMTSLNLY 579
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 579
QY 228 IRLTYNTREGK-----VDQYGE-----WRRKLSMMLTDKENPVSELLERGVNH- 271
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 271
QY 580 PAGDPRLAQAQLWEAEFLAEAFQRRMAGMQVFTFAERSLEDEINRTAEDL-PIPAT 638
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 638
QY 272 -----NMTLPEGLQD-----VADTLAPKFGV 292
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 292
QY 639 SYIVIFLY--ISLALGSYSSMSRVMD--SKATLGLGVAVVLGAVNMAAGFFSYLGI 692
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 692
QY 293 TCALFTFSFLSVLVRKHT--GVMPDWVRKPLVAAGLTPINATVTSFGLILWCGF 351
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 351
QY 693 RSSLVLQVPPVLVLSVGDADNIFIFVLEYQRLPRRPGEPREVIHGRALGRVAPSMLLCSL 752
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 752
QY 352 LYN-ALVNVSPFLLICIGIDDLFIMCAEWHR--TNPQHSPEKRIKIGTLSEAAVAISITSL 408
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 408
QY 753 SEACIFLGAITPMPAVRTALTSLGLAVLDFLLQ-----MSAFVALL----- 795
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 795
QY 409 TDIATPAMGCTTLPQVGFQCFVYQIIIFLGPVLAAYAAEQNGQHVLLIR 468
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 468
QY 796 -----SLDSKROEASRLDVCCCKVQPELPP----- 820
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 820
QY 469 KAVDPKTESPVKLWLLSGVNRQDEARRN-----SRKVPVEKEGKSKFEVVE 522
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 522
QY 821 -----PGQG--EGLLGFQKAYAPFLHMITRGVVVLLFLALFGVLSYSMCH 866
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 866
QY 523 KLEHTLEKDDDDPGHNSSEETLVSKVFREIIGPILQSKTQVCALLILYLVISLAIGGLN 582
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 582
QY 867 ISVGLQDELALPKDSYLLDYFLNLR-YFEVGPVYFV-----TTLGYNFSSEA 914
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 914
QY 583 IKEGLDPKLLVRESFYLKSFYEIIDETFWREGLOMQVWVNNPPDLTFPTTRKGF----- 637
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 637
QY 915 GMAIACSSAGCANNFSFTQKIQYATEFPEQSYLAIPASSWVDDPIDWLTPSSCCRLYISGP 974
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 974

Db 638 -----EMMAEFENTQYTMHPNATMI-----WL-----RAY----- 662
QY 975 NKDKFCPSTVNSLNCNKNCSITWGSVPSPVEQPHKYLPHFLNDRPNKCPKGGLAAYST 1034
Db 663 --EHHLETEVHELNIK-----PNSVTWYTRCDMLI-----VAGRRRLWQM 703
QY 1035 SV---NLTSQ--GOVLASRFMAVHKPKNSQDYTEALRAAELAANITADLRKVPGTDP 1089
Db 704 DMWANTYTFETPRITAFRQLGRNRYTPTDHTHSCKLMESIA-----DK 749
QY 1090 FEVEPYTITNVFY---EQYLTLPEGLF--WLS-LCLVPTFAVSCLLGLDLRGLMLLL 1143
Db 750 YSMENVTTTFHEYYPFADQYLELT-SLFQNMISDLCTI--FAVSWWMP-ELICAVAIVL 805
QY 1144 SIWMLVDYTGFMALWDISYNAVSLINLVSAVGNSEVFVSHITSPALSTKPTWLERAKE 1203
Db 806 SIASINVGVLGFMFSGVGNLDSVSIITVICIGFSVDLSAHIAVAFSQSYGNSH-ARAVA 864
QY 1204 ATISGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFFRLLNLLITLLGLLHGLVFLPVLIS 1263
Db 865 ALETLGHVPVFLGASSTVL-GILLTLVDSYVQIFFKTVFLVIN-PSILHGLIFLPIFLM 922
QY 1264 YVGPDVNPALALEQKRAEAV 1284
Db 923 KVVREVKE-----ESPRDDQSV 939

RESULT 99
Q60W90 CAEBR PRELIMINARY; PRT; 951 AA.
AC Q60W90;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG19221.
GN Name=CBG19221;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAC01000098; CAE72125.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0007275; F:development; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR006162; Ppantne.S.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; I.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
DR PROSITE; PS50156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 951 AA; 108848 MW; 46D9454B91000ED4 CRC64;

Query Match 6.1%; Score 424; DB 2; Length 951;
Best Local Similarity 20.5%; Pred. No. 7.6e-21;
Matches 220; Conservative 177; Mismatches 367; Indels 308; Gaps 42;

QY 339 QFFQGWGTWASWPLTILVLSVPPVLAAGLVFTE--LTTDPVLEWSPNQARSEKAP 396
Db 10 EYFVNVGVVHKFRLLYLLSPAILTFAFGLWFEKQTTDRDPQVFPSPANAPWRYERAV 69
QY 397 HDQH-----FGP-----FFRTNQVILTAPNRSSRYSDLSLLGPKNFSGILDLDLLELLE 446
Db 70 LTEHWPDLDEKFPWGSYDLHGVDVIASGRVHPDY-----GRPN-----ILNIRYLDEVAR 121
```



```
Qy 808 DV-CCCV-----XPQLPPP-----CQEGEL----- 827
Db 535 DVLCCCLPALPSTTHDLNTORSYHYHOQPOQOQOQOQNNIEAQRHEBETLIECQSKSC 594
Qy 828 ----LLGFFQKAYAPFLHMITRGVW----LLLFLALFGVSLYSMCHISVGLDQELALPK 879
Db 595 LSFSLSRFAIKHYAPF-----ITRSVKVFGMLLSGILGVSLFASMKLPDGLELTDLPVQ 650
Qy 880 DSYLLDYFLNRYFEVGAPVY-----FVTTLGYNFSSEAGMNAICSSAGCNPFSF--TQ 932
Db 651 NTN-----EHRFLSVQGLFGFYSPFAVTQG-----DPEYENNQ 684
Qy 933 KI--QYATEFPEQSVL-----AIPASSWDDDFIDWLTPSSCC-----RLYI 971
Db 685 KLLHEYHEAFVRVSHVIKNDNGGLP--DFWLSLFRDVLNQLQAFRDYREGRIITQERWYS 743
Qy 972 SGPNK-----DKFCPSTVSNLNLKNCMSITWGSVRPSVEQFHKYL- 1012
Db 744 NASNDAILAYKLLVQTGHVDNPIDKSLVTQVRLVD-----SEGVINP--EAFYNYS 793
Qy 1013 PWFNLNDRPNIKCPKGLAAYSTSVNLTSQGQVLASRFMAYHKPLKNSQDYTEALRAAREL 1072
Db 794 AWAMND-----VLAYG-----ASQLPFYLHGLSDTADIKTMISQIREL 831
Qy 1073 AANITADLRKVPCTDPAFEVEPYTITNVFVYEQYLTLPEGLFMLSCLVPTFAVSCLLG 1132
Db 832 CSRF--ESRGLPN-----YPSGIPPIFWEQYMLRGLLKAYGCALLAVFVVSLLT- 881
Qy 1133 LDLRSGLLNLLSIVMILVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHI 1185
Db 882 LSGWAALLIVNLMMQIQLLGVMLLLGKLSAIPAVILIASIGLGVGIVTHV 934
```

Search completed: April 11, 2006, 00:45:18
Job time : 316 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:45:32 ; Search time 53 Seconds
(without alignments)
2077.811 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2403	34.8	1318	2	US-09-949-016-10152
2	2402.5	34.8	1278	2	US-09-462-136-2
3	2385	34.5	1319	2	US-09-462-136-4
4	1329	19.2	1170	2	US-09-462-136-6
5	1078	15.6	1296	2	US-09-462-136-9
6	651.5	9.4	1447	1	US-08-540-406-19
7	651.5	9.4	1447	1	US-08-656-055-19
8	651.5	9.4	1447	2	US-08-954-668-19
9	651.5	9.4	1447	2	US-08-268-140-5
10	651.5	9.4	1447	2	US-08-918-658-19
11	651.5	9.4	1447	2	US-09-724-631-19
12	651.5	9.4	1447	2	US-08-954-701A-19
13	651.5	9.4	1447	2	US-09-807-007-6
14	651.5	9.4	1447	2	US-09-754-032-19
15	651.5	9.4	1447	2	US-08-916-140-19
16	651.5	9.4	1447	4	PCT-US95-13233-19
17	646.5	9.4	1434	1	US-08-540-406-10
18	646.5	9.4	1434	2	US-08-656-055-10
19	646.5	9.4	1434	2	US-08-954-668-10
20	646.5	9.4	1434	2	US-08-918-658-10
21	646.5	9.4	1434	2	US-09-724-631-10
22	646.5	9.4	1434	2	US-08-954-701A-10
23	646.5	9.4	1434	2	US-09-754-032-10
24	646.5	9.4	1434	2	US-08-916-140-10
25	646.5	9.4	1434	4	PCT-US95-13233-10
26	639.5	9.3	642	2	US-09-248-796A-16161
27	628	9.1	1296	2	US-08-857-636-60

28	584	8.5	1203	2	US-09-293-505-2	Sequence 2, Appli
29	584	8.5	1203	2	US-09-060-939A-2	Sequence 2, Appli
30	583.5	8.4	1182	2	US-09-293-505-7	Sequence 7, Appli
31	583.5	8.4	1182	2	US-09-060-939A-7	Sequence 7, Appli
32	581.5	8.4	531	2	US-09-248-796A-16160	Sequence 16160, A
33	580	8.4	1203	2	US-09-207-857-2	Sequence 2, Appli
34	580	8.4	1203	2	US-09-909-280A-2	Sequence 2, Appli
35	579	8.4	1146	2	US-09-807-007-1	Sequence 1, Appli
36	526.5	7.6	1311	1	US-08-540-406-4	Sequence 4, Appli
37	526.5	7.6	1311	1	US-08-656-055-4	Sequence 4, Appli
38	526.5	7.6	1311	2	US-08-954-668-4	Sequence 4, Appli
39	526.5	7.6	1311	2	US-08-918-658-4	Sequence 4, Appli
40	526.5	7.6	1311	2	US-09-724-631-4	Sequence 4, Appli
41	526.5	7.6	1311	2	US-08-954-701A-4	Sequence 4, Appli
42	526.5	7.6	1311	2	US-09-754-032-4	Sequence 4, Appli
43	526.5	7.6	1311	2	US-08-916-140-4	Sequence 4, Appli
44	526.5	7.6	1311	4	PCT-US95-13233-4	Sequence 4, Appli
45	474	6.9	1286	2	US-09-268-140-3	Sequence 3, Appli
46	464.5	6.7	1299	2	US-08-460-900C-62	Sequence 62, Appli
47	464.5	6.7	1299	2	US-08-674-509B-48	Sequence 48, Appli
48	464.5	6.7	1299	2	US-08-954-698-48	Sequence 48, Appli
49	464.5	6.7	1299	2	US-09-639-695-62	Sequence 62, Appli
50	464.5	6.7	1299	2	US-09-448-188-48	Sequence 48, Appli
51	464.5	6.7	1299	2	US-08-954-128-48	Sequence 48, Appli
52	464.5	6.7	1299	2	US-08-954-740-48	Sequence 48, Appli
53	438.5	6.3	1285	1	US-08-540-406-6	Sequence 6, Appli
54	438.5	6.3	1285	1	US-08-656-055-6	Sequence 6, Appli
55	438.5	6.3	1285	2	US-08-954-668-6	Sequence 6, Appli
56	438.5	6.3	1285	2	US-08-918-658-6	Sequence 6, Appli
57	438.5	6.3	1285	2	US-09-724-631-6	Sequence 6, Appli
58	438.5	6.3	1285	2	US-08-954-701A-6	Sequence 6, Appli
59	438.5	6.3	1285	2	US-09-754-032-6	Sequence 6, Appli
60	438.5	6.3	1285	2	US-08-916-140-6	Sequence 6, Appli
61	438.5	6.3	1285	4	PCT-US95-13233-6	Sequence 6, Appli
62	254.5	3.7	529	2	US-09-500-063-2	Sequence 2, Appli
63	151	2.2	232	2	US-09-270-767-31821	Sequence 31821, A
64	150	2.2	1237	2	US-09-332-522E-6	Sequence 6, Appli
65	145	2.1	1467	2	US-09-252-991A-17657	Sequence 17657, A
66	140.5	2.0	1051	2	US-09-252-991A-16989	Sequence 16989, A
67	140	2.0	1005	2	US-09-328-352-4877	Sequence 4877, Ap
68	138	2.0	826	2	US-09-902-540-12490	Sequence 12490, A
69	137.5	2.0	1040	2	US-09-489-039A-10700	Sequence 10700, A
70	136.5	2.0	1032	2	US-09-332-522E-94	Sequence 94, Appli
71	133	1.9	117	2	US-09-248-796A-16571	Sequence 16571, A
72	130.5	1.9	792	2	US-10-055-364-42	Sequence 42, Appli
73	130.5	1.9	808	2	US-08-804-439A-14	Sequence 14, Appli
74	130.5	1.9	808	2	US-08-720-229-14	Sequence 14, Appli
75	130.5	1.9	1069	2	US-09-252-991A-24533	Sequence 24533, A
76	127.5	1.8	1607	2	US-09-902-540-16765	Sequence 16765, A
77	126	1.8	801	2	US-09-710-279-2020	Sequence 2020, Ap
78	126	1.8	808	2	US-09-134-001C-3105	Sequence 3105, Ap
79	125.5	1.8	1123	2	US-09-252-991A-27976	Sequence 27976, A
80	123.5	1.8	789	2	US-09-252-991A-28767	Sequence 28767, A
81	123.5	1.8	1065	2	US-09-328-352-8064	Sequence 8064, Ap
82	123	1.8	1033	2	US-09-252-991A-21000	Sequence 21000, A
83	123	1.8	1049	2	US-09-489-039A-13014	Sequence 13014, A
84	122.5	1.8	1052	2	US-09-252-991A-30591	Sequence 30591, A
85	120	1.7	863	2	US-09-902-540-10725	Sequence 10725, A
86	118.5	1.7	735	2	US-09-902-540-14015	Sequence 14015, A
87	118.5	1.7	1033	2	US-09-252-991A-20611	Sequence 20611, A
88	115.5	1.7	914	2	US-10-104-047-2923	Sequence 2923, Ap
89	115	1.7	651	2	US-09-902-540-11066	Sequence 11066, A
90	115	1.7	810	2	US-09-902-540-12414	Sequence 12414, A
91	114	1.7	976	2	US-08-311-731A-363	Sequence 363, App
92	114	1.7	1019	2	US-09-252-991A-24417	Sequence 24417, A
93	112.5	1.6	697	2	US-09-603-208A-226	Sequence 226, App
94	112	1.6	517	2	US-09-252-991A-21762	Sequence 21762, A
95	112	1.6	750	2	US-09-902-540-11016	Sequence 11016, A
96	112	1.6	1051	2	US-09-543-681A-8143	Sequence 8143, Ap
97	111.5	1.6	506	2	US-09-540-236-2360	Sequence 2360, Ap
98	111	1.6	1025	2	US-09-711-164-443	Sequence 443, App
99	111	1.6	1025	2	US-09-492-709A-283	Sequence 283, App
100	111	1.6	1048	2	US-09-489-039A-7216	Sequence 7216, Ap

ALIGNMENTS

```
RESULT 1
US-09-949-016-10152
; Sequence 10152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10152
; LENGTH: 1318
; TYPE: PRN
; ORGANISM: Human
US-09-949-016-10152

Query Match          34.8%; Score 2403; DB 2; Length 1318;
Best Local Similarity 39.7%; Pred. No. 3.9e-218; Indels 110; Gaps 26;
Matches 525; Conservative 238; Mismatches 451;

QY  2  ABAGLRGMLLWA-----LLRLAQSEPYTHIQFGYCAFYDECG-----KNPELSG 47
DB  32  ASALRGHSMTARGALGLLLLLCPAQVFSQ-----SCVWTEGCGIAYGDKRYNCEYSG 86
QY  48  SLMTLSNVSCLSNTPARKITGHLILLQKICRLTYGPNQACCSAKQLVLSLEASLSITK 107
DB  87  -----PPKPLPKDGYDLVQLCPGFPGF-NVSLCCDVRQLTKLQNLQPL 131
QY  108  ALLTRCPAGSDNFVNLHCHNTCSNQSLFINVTR-----VAOLGAGOLPAVAYEAFYQHS 163
DB  132  QFLSPCPCFYNLNLLFCBLTSCPSQSQFLANTATEDYDPTNQTNTNKLQYVVGOS 191
QY  164  FAEQSYDSCSRVRPAAATLAVGTWCGVYGSALCNAQRWLNFGQDGTGNGLAPLDTIT---- 219
DB  192  FANAMYNACRDVEAPSSNDKALGLLCKGDADA-CNATNWEYMNKNDGQAPFTITPVFS 250
QY  220  ----FHLLEPGQAVGSIQPLNKGVARCNESQGGDVATCSCQCAASCPIARPQ----- 270
DB  251  DFPVH-----GMEPMNNAATKGCDSVDEVTAPCSCDCSIVCGPKPQPPPPAPW 300
QY  271  ----ALDSTFVLGMPGSLVLIILCSVEAVVITLL-----VGPRVAPARDKS 314
DB  301  TILGLDAMVIMWITWMAELLVFFGAFFAVCWKRKYFVSEYTPDSDNIAFSV-NASDKG 359
QY  315  KMVDPKKGTSLSKLSFSTHTLILGQFFQGWGTWVASWPLTILVSLVPIVVALAAGLVFTE 374
DB  360  E-----ASCCDPVSAAFEGCLRLFTRWGFCVNRNPGCVIFFSFLVITACSSGLVFR 412
QY  375  LTTDPVELMSAPNSQARSKAFHQHGFPGFPFRNTQNVILTAPNRSSVRYDSLLGPK-NFS 433
DB  413  VTTNPVDLMSAPSSQARLEKEYFDQHFPGFFRTEQLIIRAPLTDKHIQYFYPGSDVDPFG 472
QY  434  GILDLLLELLELLEQLRHQLQVWSPQAQRNLSLODICVAPLNPDNTSLDYCCINSLLQY 493
DB  473  PPLDIQLHQVLDLQIAIEN--ITASYDNETVTLQDICIAPLSPYNT---NCTILSVLNY 527
QY  494  FQNNRTLLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALSCMADYAGAPVFPF 553
DB  528  FQNSHSLVDHKKGDDEF--VYADYHTHFLYCVRAPASLNDTSLLDHPCLTGTFGGVFPW 584
```

```
QY  554  LAIGGYKGKDYSEAEALIMTFSLNYPAGDPRLAQAALWBEAFLEEMRAFORRMAGMFQV 613
DB  585  LVLGYYDDQNNYNATALVITFPVNNYNDTEKLOQAQWEKEFINFKYKN---PNLTI 641
QY  614  TFTAERLEDEINRTAEDLPFIPATSYIVIFLYISLALGSYSSWSRWSKATILGCGV 673
DB  642  SFTAERSTEDLNRESDSVFTVVVISYAIMPLYSIALGHIKSCRRLLLVDSKVSIGIAGI 701
QY  674  AVTLGAYMAAGPESYLGIRSSILVLOVPLVLSVGADNIFIFVLEYQRLPRRPGPRE 733
DB  702  LIVLSSVACSLGVPSYIGLPLTIVIEVIFPLVAVGVDNIFILVQAYQDERLQGETLD 761
QY  734  VHGALGRVAPSMMLLCSLSEACIFFLGALTPPAVRTFALTGSLAVILDFLLQMSAFVA 793
DB  762  QQLGRVLGEVAPSMFLSPSETVAFVLGALSVMAVHTFSLFAGLAVFIDFLQLITCFVS 821
QY  794  LLSLDSKRQREASRLDVCCCVKPBELPPGQ-GEGLLLGFPQKAYAPFLLHWHITRGVLL 852
DB  822  LLGLDIKRQEKRLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPILLKDDMRPIVIAI 881
QY  853  FLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFS 912
DB  882  FVGLSFSIAVLNKVDIGLQSLSPDDSYMVDYFKSISQYLHAGPPVYFVLEEGHDYIS 941
QY  913  EAGNNAICSSAGCCKNFSTQKIQYATFPPEQSYLAIPASSWVDDPFDWLTTP-SSCCRLYI 971
DB  942  SKQNMVCGGCGCNDLSLQQIFNAQLDNVTRIGFAPSSWIDDYFDWVKPQSSCCRV-- 999
QY  972  SGPNKDKPCPSTVNSLNLCKNCSIT-MGSVVRPSVEQPHKYLPMFLNDRPNIKCPKGGLA 1030
DB  1000  -DNITDQFCNASVDDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHA 1057
QY  1031  AYSTSVN--LTSQDVLASREMAVHKPLKNSQDYTEALRAARELAANITADLRKVPGETP 1088
DB  1058  AYSSAVNILLGHGRVGTGYFTMTYHTVLTQTSADFDALKARLTASNVT-ETMGINGS-- 1114
QY  1089  AFEVPPYPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLNLLSIYMI 1148
DB  1115  AYRFPFSVFFVFEQYLTIDDDIFNLGVSLGAFLVTVWLLGCELWSAVIMCATIAMV 1174
QY  1149  LVDTVGFMAWDISYNAVSLNLSAVGMSVEFVSHITRSFAISKPTWLERAKAETISM 1208
DB  1175  LVNPMGVNMLWGISLNAVSLNLMSCGISVEFCSHITRAFTVSMKGSRVERAEALAHM 1234
QY  1209  GSAPVAGVAMTNLPGILVLGLAKAQLIOIFFRNLNLLITLLGLHGLVFLPVLISVSGD 1268
DB  1235  GSSVFSGITLTKFGGIVVLAFAKSQIFQIFFRMYLAMVLLGTHGLIFLPLVLLSYGPS 1294
QY  1269  VNPA 1272
DB  1295  VNKA 1298

RESULT 2
US-09-462-136-2
; Sequence 2, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: PRN
```

```

; ORGANISM: Homo sapiens
; US-09-462-136-2

Query Match      34.8%; Score 2402.5; DB 2; Length 1278;
Best Local Similarity 39.8%; Pred. No. 4.2e-218;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLRAQSEPTTHIQPGYCAFYDRCG-----KNPELSGSLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCWYGCAGIAYGDKRYNCBSYG-----46

QY 60 NTPARKITGDHLLILQKICPRLTYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLQVELCPGFFG--NVLCCDVRQLQTLKONLQLPQLFSLRCSFCFN 103

QY 120 FVNLCHNTCSNQSLFINVTR----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
DB 104 LNLFLCELTCSPQSOFLNVTATEDYDVPVTQTKNVKELQYVYQSPFANNAYNACRDV 163

QY 176 RYPAATAATLVGTCMGVYGALCNARWLNFQDGTGNGLAPLDIT-----PHLLPQGA 228
DB 164 EAPSNDRKALGLLCKGDADA--CNATNWIEYMFNKDNGQAPFTITPVFSDFPVH-----215

QY 229 VSGGIQPLNEGVARCNESQDDVATCSODCAASCPAIAPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGDSESDVETAPCSQDCSIVCGPKPQPPPPPPAPMTILGLDAMYIM 272

QY 280 QMPSGLVLIILCSFVAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLVPGFAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

QY 327 DKLSSTHTLLGQFFGWTGTVASWPLTILVLSVPVVALAAGLVFTLTTPDVLWAP 386
DB 325 DPVSAAFEGCLRLRTRWGSFCVPGFCVIFPSLVEITACSSGLVFRVVTNPVFLWAP 384

QY 387 NSQARSEKAFDHOEGPFRITNOVILTAPNRSRYDSLLGPK-NFSGILDLDLLELL 445
DB 385 SQQARLEKEYFOHGGPFRTEQLIRAPLTDKHIYQYPVPGADVPFGPPLDIQLHQVL 444

QY 446 ELQERLRHLQWSPQAQRNLSIQDICYPALNPDTSLYDCCINSLLQYFQNNRTLLTLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDILCLAPLSYNT---NCTILSVNLFQNSHVDHKK 499

QY 506 NOTLMGQTSQVMDKHDFLYCANAPITFKDGTALALSCMADYGAPEVFLAIGYKGDYS 565
DB 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLDHPCLGTGFGVPFVPLVLGGYDDQNV 556

QY 566 EBAALIMTSLNYPAGDRLAQAQKLEAEFLSEMRAFQRRMAGMFQVFTFAERSLBRI 625
DB 557 NATALVITFPVNNYNDTEKQRAQAWEKEFTNFVKNYK---PNLTISFTAESRIDEEL 613

QY 626 NRTTAEDLPIPATSVIVIFLYISLALGSYSMSRVMDSKATLGLGGVAVVLGAVMAAG 685
DB 614 NRESDSVFTVVISVAINFLYISLALGHKSCRLLVDKSVLSLGTAGLILVSSVACSLG 673

QY 686 FFSYLGIRSSVLQVFPFLVLSGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
DB 674 VFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAYQDRERLOGETLQOGLGRVLGEVAP 733

QY 746 SMLLCSLEAICFFLGALTPPAVTPALTSLGLAVILDFLOMSAFVALLSLSDRQAS 805
DB 734 SMFLSSFSFETVAFFFGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLGIDIKRQEN 793

QY 806 RLDDVCCVKPOBLPPPGQ--GEGLLIGFFQKAVAPPLLHWITRGVVLFLALFGVLSYM 864
DB 794 RLDFICCVRGAEADGTSVQASECLFRFFNFNSPLLLKDWMPVIAIFVGLVLSIAVL 853

QY 865 CHISVGLDQELAPKDSYLLDYFLFLNRYFEVGAPEVFTTILGYNFSSAGNAATCSSAG 924
DB 854 NKVDIGLQSLMPDSDSYWDYFKISQYLHAGPPVYFVLEBGHDYTSKQNMVCGMG 913

QY 925 CNNSFTQKIQVATEFPFQSYLAIPASSWVDDFIDWLTP--SSCCRLYISGPNKKPCPST 983
DB 925 CNNSFTQKIQVATEFPFQSYLAIPASSWVDDFIDWLTP--SSCCRLYISGPNKKPCPST 983

```

RESULT 3

```

US-09-462-136-4
; Sequence 4, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1319
; TYPE: PRT
; ORGANISM: Mus sp.
; US-09-462-136-4

Query Match      34.5%; Score 2385; DB 2; Length 1319;
Best Local Similarity 40.0%; Pred. No. 2e-216;
Matches 524; Conservative 231; Mismatches 440; Indels 116; Gaps 28;

QY 14 LLLRLAQSEPTTHIQPGYCAFYDRCGKNPELSGSLMTLSNVCSLSNTPKITGDHLL 73
DB 53 LLLLLCPAQVFSQ-----SCWYGCAGIAYGDKRYNCBSYG-----46 100

QY 74 LQKICPRLTYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNVLHCHNTCSNQ 133
DB 101 VQELCPGLFF--DNVSLCCDIQOQLTKNLQLPQLFSLRCSFCFYNLMTLFCELTCSPHQ 159

QY 134 SLFNTVTRVAQLGAGQLPA---VYAEAFYQHSFAEQSYDSCSRVYPAATLAVGTCM 189
DB 160 SQFLNVTATEDYFDPKTPENTKNVKELEYVYQSPFANNAYNACRDVREAPSNEKALGLIC 219

QY 190 GYVSALCNARWLNFQDGTGNGLAPLDI-----TFHLLPEQAVGSGIQLNEGVARCN 244
DB 220 GRDARA--CNATNWIEYMFNKDNGQAPFTITPVFSDSLIL-----GMEPMNATKGCN 270

QY 245 ESQGDVATCSQDCQCAASCPAIAEQALDSTFYLGQMPGSLVLIILCSFVAVVTLLVG 304
DB 271 ESDVETGPCSCQDCSIVCG--PKQILQIGGGGGG--GLDAMYVIMVTVYVFLVFPFG 327

QY 305 FRVA-----PARDKSNWDPKKGTSLSDKLSSTHTLLGQF 340
DB 328 ALLAVWCHRRRYFVSEYTPIDSNIAFSVNSSDKGE-----ASCCDPLGNAFDCLRRM 380

```



```
QY 822 GQEGLLGFFQKAYAPFLLHWTITRGVLLLLFLALFGVSLYSMCHISVGLDQSLALPKDS 881
Db 729 -GNEETKESFLKTFYFKMLTQ--KRLIIIIISAWFTSLVFLPEIQFGLDQTLAVPQDS 784
QY 882 YLLDVPFLNRRYFEVGCAPVYVTTILGYNFSSSEAGNAICSS-AGCNPFSTQKIQVATEF 940
Db 785 YLVDYPRKDYSELNVPVYVVK-NLDTKRNQOQKICGKFTTCERDSLAVLE---Q8 840
QY 941 PRQSYLAIPASSVDDDFIDWLTSS--CCRLYISGNPKDKFCPSTVNSLCLKNCMSITM 998
Db 841 RHRSTITEPLANWLDYFMLFNPQNDQCEL---KGTDEVCPSPSPSRC-ETCFQ--Q 894
QY 999 GSVR-----PSVEQPHKYLWPLNDRPNKCPKGLAAAYSTSVNLTSDGQVLASRFMAY 1052
Db 895 GSWNTNMSGFPKGPMEYLSWIN-APSDPCPLGGRAPYSTAL-VYNETSASVFRTA 952
QY 1053 HKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPVFPVYTTNVFYEQYLILPRG 1112
Db 953 HHPLRSQKDFIQAY-----SDGVRISSPPELDMFAYSPPYIFVQYQTLGPUT 1001
QY 1113 LFMLSLCLVPTFAVSCLLGLDLRSGLNLLGLVLMILVDTVGFMALWDISYNASVLINLV 1172
Db 1002 LKLGSAIILIPFISVFL-QNIRSFLALVVTMIIIVDIGMALLGLISLNAVSLVNDI 1060
QY 1173 SAVGMSVEFVSHITRSFAI---STKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGL 1229
Db 1061 ICVGLGVPCVHIVRSFTVVPSETKKDANSRVLYSLNTIGESVIGKITLTKFIVGCVLAF 1120
QY 1230 AXAQILQIPFRLNLLITLLGLHLGLVFLPVLSVYG 1266
Db 1121 AQSKIFDVFPYRMWFTLIIAVALHALLFLPALLSLFG 1157

RESULT 5
US-09-462-136-9
; Sequence 9, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstea, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1296
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-462-136-9

Query Match
Best Local Similarity 24.0%; Score 1078; DB 2; Length 1296;
Matches 321; Conservative 237; Mismatches 492; Indels 288; Gaps 41;

QY 76 KICPLRYTGPNTQACCSAKQLVSLBASLITKALITRCPACSDNFVNLHCHNTCSPNQSL 135
Db 63 BFCPHLLTGDN-KLCTPQSAEGLTKQIAQARHILGRCPSCFDNFAKLWCBEFTCSPNQD 121
QY 136 FINVTRVAQL--GAGQLFAVAYEAF-----YQHS--FABQSYDSCSRVPAATAVLAVG 186
Db 122 FVSISEMPKIEKGEFTPEYQPAEAYVNTVEYRLSTDFAEQMFSSCKDVTFGGQPALRV- 180
QY 187 TWCYGVYGCALCNAQRLWLFQDGTGNL-APLDITFHLLERQAVQSGIQP-LNEGVARCN 244
Db 181 -MC-----TSTPCPLTNLWLFPIGTQNLIDNIPHTKFLDYDPKTPPSDRSTYNNVFTGCD 236
QY 245 ESQGDVATCSQDCAACPAIARQALDSTPYLQM-----PGSLVLIILICSVFAVVT 299
```

```
Db 237 KSARVGMFACSTSEC--NKREYANLIDLDGKTSQTCNVHGIACLNIFVMLAFISLAV 294
QY 300 ILLVGF-----RVAPARDKSKWVDPKKGTSLSKLSFSFTHLLGQFP 341
Db 295 LLCVGFVFTSYDEVDYTNLRQTOSGESPKRNRK-----RTGAWI-----HNFMENNA 342
QY 342 QCWGVTWASWPLTILVLSVIPVVALAAGLVFTTETDTPVELWSPANSQARSKAPFHDHF 401
Db 343 RDIGMMAGRNPKSHFFIGCAVLIFCLPGMIYHKESTNVVDMSSPRSRARQEWVFNANP 402
QY 402 GPFTTNQVILTAPNRSYRYDSLLGPKNFGSGLDLDLLELLELLEQLERHLQWSPBA 461
Db 403 GRPQYQOITMLL--SHRDFQSSGKLYGP-----VFHKDIFEBLFDILNAIKNISTQDSG 455
QY 462 QNRSIQDICVAPLNPDTNLSLVCINSILQVFNQNRITLLLTANQT-----L 509
Db 456 -RTITLDDVCYRPMGPG---YDCLIMSPNTYFQGNKEHLDKMSKBEVTSBDDDAFYF 510
QY 510 MGQTSQVDWKDHFYLCANAPLTFKGTALALSCMADYGAAPVFPFLAIGYKGYKDYSEARA 569
Db 511 SSEATTDWMNHMAACIDQMSQK--TKSGLSCMGTYGSPSAPNM-VFGKSTNHOANS 567
QY 570 LIMTSLNYPAGDPRLLAOKLWREAFLEEMRAFORMAGMEQVTFTAERSLEDEINRTT 629
Db 568 IMMTILVTQ--RTEPEIQKAELEKEFLKCKEYREKSPKVI-PSFMAERSITDEIENDA 624
QY 630 AEDLPFATSYIVIFLYISLALGSY-----SSWSRVWVDSKATILGLGGVAVVLGVAMAAM 684
Db 625 KDEIVTVVIALAFLIGYVTFSLGRYFVCENQLWS-ILVHSR----- 664
QY 685 GFPSYLGIRSSILVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHVHIGRALGRVA 744
Db 665 GF----- 666
QY 745 PSMLCSLSEACFFFLGALTMPAVRTPALTSGLAVILDFLQMSAFVALLSLDSKROBA 804
Db 667 -----TDLPAIRTFCLVAGLAVLIDVVLHCTIFLALFVMDTQRELN 707
QY 805 SRLDVCCCKPQELPPQCGEGL-----LLG-----PFQKAYAPFLHWTITRGV 849
Db 708 G-----KP-BFFPYQIKDLLGAYLIGRQATDTFTMQTFHFQVAFPLHMRTRIT 758
QY 850 LLLFLALFGVSLYSMCHISVGLDQSLALPKDSDLYDLFLNRYFEVGCAPVYVTTILGN 909
Db 759 GIIFIASITVTVLSSKISVGFQSMATEKSYISTHFRYLDKFPDVGPPVFTVDGELD 818
QY 910 FSSEAGMNAICSSAGCNPFSTQKIQVATEPPEQSYLAIPASSWVDDFIDWLT-PSCCR 968
Db 819 WHRPDVQNKFCCTFPGCSDTSPGNIMNYAVGHTEQYLSGEMVYNDVLEWISRSKSPCK 878
QY 969 LYISGPNKDKPCPSTVNSL-----NCLKNCSITMGSVRPSVEQFHKYLPWFLNDRPNKC 1024
Db 879 VYVHDPN--TFCSTNRNKSALDDKACRTCMDF----- 908
QY 1025 PKGGLAAYSTSVNLTSDGQVLASRFMAYVKPL--KNSODYTEALRAARELANITADLRK 1082
Db 909 --DGEASFDAISFTSRGRIQASQWTFHKLSINSDDPIKAMDTAMVSRRLRSI-- 964
QY 1083 VPGTDPAFEVFPYTTINVFYEQYLITLPEGLFMLSCLCLVPTFAVSCLLGLDLRGLNML 1142
Db 965 ---DDTAHVAYSKIPFPYEQYSTIMPIITQTLPITVVGFGIICVTLGDVKAACAV 1020
QY 1143 LSIWMLVDTVGFMALWDISYNASVLINLVSAVMSVEFVSHITRSFAISTKPTWLERAK 1202
Db 1021 ICQVS-----NYPHIVSS-GILIEFSVNLKGYACSLRQAKRAE 1060
QY 1203 EATISMGSAVFAGVAMTNLPGILVGLAKAQLIQIFPRLNLLITLLGLHLGLVLPVL 1262
Db 1061 STVSGIPIILUSGPVVTWAGSTMFLSGAHLQIITVFPKGLFITITVSAVHALIITLPI 1120
QY 1263 SYVGPDVNPALAEQKRAEBAVAAVMVASCNHPSRVSTADNIYVN-----HSFEGS 1314
```

Db	1121	AFGSRGHSSETNDNDQDACVLS--PTAESHSINVEGILNRPSSLDDASHILDPL	1178
Qy	1315	IKGAGAINFLPNNGROF	1332
Db	1179	LKAEGGIDKAI---GRDF	1193
RESULT 6			
US-08-540-406-19			
; Sequence 19, Application US/08540406			
; Patent No. 5837538			
; GENERAL INFORMATION:			
; APPLICANT: SCOTT, MATHEW P			
; APPLICANT: GOODRICH, LISA V			
; APPLICANT: JOHNSON, RONALD L			
; TITLE OF INVENTION: Patched Genes and their Use			
; NUMBER OF SEQUENCES: 19			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert			
; STREET: Four Embarcadero Center, Suite 3400			
; CITY: San Francisco			
; STATE: CA			
; COUNTRY: US			
; ZIP: 94111			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/540,406			
; FILING DATE: 06-OCT-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Rowland, Bertram I			
; REGISTRATION NUMBER: 20015			
; REFERENCE/DOCKET NUMBER: a60190-1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415-781-1989			
; TELEFAX: 415-398-3249			
; INFORMATION FOR SEQ ID NO: 19:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1447 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-540-406-19			
Query Match 9.4%; Score 651.5; DB 1; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVGSGIOPLNEGVARCNESQDDVATCSQDCAASCPAIARPOALDSTFYLGQMPGS	284
Db	24	PGPAGGRRRTTGGLRAAAPRD-----YLRHPSYCDAAFALEQI---65	
Qy	285	LVLIIILCSFVAVVTILLVGRVAPARDKSMVDPK-----KGTSLSDKLSFSTHTLL	337
Db	66	-----SKGKATGRKAPLMRAKFORLLFLKLCYIQKNC	98
Qy	338	GQFPQGGTGWASWPLTILVLIPVVALAAGLVFTLTDPVELKSAPNQSARSEKAPH	397
Db	99	GKF-----LVVGLLIIFGAFVAGLKAANLETNVELWVEVGGRVSRRELYNT	143
Qy	398	DQHFQPPFRTN-QVILTAPNRSSRYSDSLIGPNKFSGIILDLLELLE--LQERLRHL	454
Db	144	RQKIGEEAMFNPLMIQTPKEG-----ANVTTEALLQHLSALQASRVHV	190
Qy	455	QWGSPEAQRNLSLODICVAP-----LNPDMTSLYDCINSLLQYFQNNRTLLLLTA	505
Db	191	YMYN----ROWKLEHLCYKSGELITETGYMDQIIIEYLYPCLIIITPDCFWEGAKLQSGTA	246

Qy	506	NOTLMGQ-----TSQVDWKDHFY-----CAN-----527	
Db	247	--YLLGKPLRWTFPLPFLBELKKNYQVDSWBEMLNKAEVGHGVMRCPCLNPADPDC	304
Qy	528	-----APL-----TFKQGT-----NALSCMA	544
Db	305	PATAPNKNSTKPLDMALVNLGGCHGHSRYKMHMQEELIVGGTYKNSTGKLVSAHALQTMF	364
Qy	545	DYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPLRAQAQKLEWEEAFLEEMRAFQ	604
Db	365	QLMTPKQMY---EHFKGYEY-----VSHINWNE---DKAAALILEAKQRYVEVVHQSV	411
Qy	605	RRMAGMFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYXISLALGSYSSSRVMDVS	664
Db	412	AQNSTQKVLSTTT-TTLDDILKSFSDSVIRVASGYLLMLAYACLTM---LRMD--CSKS	465
Qy	665	KATLGLGGVAVLGVAAAMGPFYSYLGIRSSVLIVQVPELVLSVGADNIFIFVLEYQRL	724
Db	466	QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLALGVGVDDVFLLAHAFSET	525
Qy	725	PRRGEPREVHIGBALGRVAPSMLLCSLSBAICFFLGCALTMPMPAVRTFALTSGLAVIDF	784
Db	526	GQNKRIPEDEBTGECRLKRTGASVALTISINVTAFMAALIPALPRAFSLQAQAVVVENF	585
Qy	785	LLQMSAFVALLSLSKROEASRLDVCCC-----VKQEL-----PPPG	822
Db	586	AMVLLIFPAILSMDLXRREDRLDIFCCFTSPCVSRVQVPEQAYTDTDNTRYSPPPY	645
Qy	823	QGEGL-----827	
Db	646	SSHSAFHEQTQTMOSTVOLRTEYDPHTVYVTTAEPSEISVQPVTTQDTLSCQSEST	705
Qy	828	-----LLGFFQKAYAPFLHMTITRGVLLFLFLALFGVSL	861
Db	706	SSTRDLSQFSDSSILHCLPEPCTKWTLSSFAEKHYAPFLKPKAKVVVIFLFLGLLGVSL	765
Qy	862	YSMCHISVGLDQELPKOSYLLDYFLPLNRYFEVGAPVYFVTTLGYNFSEAGNNAICS	921
Db	766	YGTTRVRDGLDITDIVPRETREYDFIAAQKFYFSF-----YNN-----803	
Qy	922	SAGCNPFSTQKIQYAT-----EFPEQSYLAIAPASS-----WYDDFIDML-----961	
Db	804	-----YIVTQKADYPNIQHLLYDLHRFSNVKYVMLEENKQLPKMWLHYFRDMLQGLQD	857
Qy	962	-----TPSS-----CCRLYISGPNKDKPCPSTVNSLNCMLKNCMSITWGS	1000
Db	858	AFDSDWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI	913
Qy	1001	VRPSVEQFHXYL--PWFLND-----RNIKCPKGLAAY--STSNNLTSDQ	1043
Db	914	INPSA--FYIYLTAWVSNDPVAYAAQANIRPREVWHDK---ADYMPETRLRIPAAEP	968
Qy	1044	VLASFMAHYKPLKNSQDYTEALRAARELANITA-DLRKVPGTDPAPEFVFYTTINVFY	1102
Db	969	IBYAQFPFVNLGRUTSDFVEAIEKVYRICSNYTSGLSSYPNG-----YFP---LFW	1018
Qy	1103	EOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAIWDLS	1162
Db	1019	EQYIGLRHWLLFISVLACTFLVCAPFLNPNWTAGII-VMVLAJMTVELFGLMGLIGIK	1077
Qy	1163	YNVSLINLVSAVMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVFAVAMTN	1220
Db	1078	LSAVPVTIIASVGIVGEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST	1133
Qy	1221	LGILVLGAKAQIQLPPFRNLNLTITLGLLHGLVFLFVILSYVG--PDVNPALALEQ-	1277
Db	1134	LIGVLMAGSEDFIVRYFFAVLAITLILGLVNLGLVLLPVLISFFGYPYEVSPANGLNL	1193
Qy	1278	--KRAEEAVALVAVSCPNHPSRVSTADNIYVNHSPFEGSIK	1317
Db	1194	PTSPPEPPPSVVRFAAMPFGHTH--SGSDSDSEYSSQTTVSG	1233

RESULT 7

US-08-656-055-19
; Sequence 19, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY 225 PQGAVGSGIQLNEGVARCNESQGDVATCSCQDCAACPAIARPOALDSTFYLGQMPGS 284
DB 24 PGRPAGGRRRTTGLRRAAAPDRD-----YLRHPSYCDAAFALEQI--- 65
QY 285 LVLIILILCSVPVAVTILLVGFVRVAPADKSKWDPK-----KGTSLSDKLSPFSTHTLL 337
DB 66 -----SKGATGRKAPLWRAKFORLLFKLGCVIQKNC 98
QY 338 GQFFQCGMTWASWPLITILVLSVIPWALAAGLVFTELTDVPVLWAPNSQARSEKAFH 397
DB 99 GRF-----LVGGLLIFGAFVGLKKAANETNVEELWVEVGGVRSRELNT 143
QY 398 DQHPGPPFRTN-QVILTAPNRSYRSDLSLLGPKNFGSGLDLDLLELE--LQRLRLH 454
DB 144 RQKIGEAAMFNQMIQTPEKG-----ANVLTTEALLQHLDLSALQSRVHV 190
QY 455 QWSPFAQRNLSLDICVAP-----LNDNTSLYDCCINSLLQYQNNRTLLTLTA 505
DB 191 YMYN-----RQWLEHLCKYSGELITETGYMDQIIEVLYPCLITPLDCFWGAKLQSGTA 246
QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
DB 247 ---YLLGKPLRTNFDPLEFLBELKKINYQVDSWBEMLNKABVGHGYMDRPRCLNPADPDC 304

RESULT 8

US-08-954-668-19
; Sequence 19, Application US/08954668
; Patent No. 6172200

QY 528 -----APL-----TFKOGT-----ALALSCMA 544
DB 305 PATAPNKNSTKPLDMALVNGCGHLSRKYMHQEBELIVGGTVKNSTGKLVSAHALQTMF 364
QY 545 DYCAPVPPPLAIGYKGDYSEABALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFQ 604
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE-----DKAAAILAEAWQTYVEVVHQS 411
QY 605 REMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGYSVSSSRVWVDS 664
DB 412 AQNSTQKVLSTFT--TLLDILKSPSDSVIRVASGYLLMLAYACLTM-----LRWD--CSKS 465
QY 665 KATLGLGGVAVVLGAVMAAMGPFYSLGIRSSILVILQVWPFVLVSVGADNIPFVLEBYQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTVLPFLALGVGVDDVFLLAHAFSET 525
QY 725 PRPGEPREHVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTFALTSLAVILDF 784
DB 526 GQNKRIFFEDRTGECLEKRTGASVALTSISNVTAFFMAALIPALRAFSLQAQAVVVFNF 585
QY 785 LIQMSAFVALLSLDKROEASRLDVCCC-----VKQOEL-----PPPG 822
DB 586 AMVLLIFAILSMNDLYRREDRLDIFCFTSPCVSRVIOVFQAYTDTHDNTRYSPPPY 645
QY 823 QGEGE----- 827
DB 646 SSHSPAHEQITMQSTVQLRTEYDPHTHVYTTABPRSEISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGPFQKAYAPFLHWHITRGVWVLLPLALPGVSL 861
DB 706 SSTRDLLSQSDSLHCLPEPCTKWTLSFAEKHYAPFLPKPKAVVVFLLGLGLGVL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRFEVGPVYFVTTGLGYSSEAGNNAICS 921
DB 766 YGTTAVRGLDLTDIVPRETREYDFIAAQFKYFSP-----YNM----- 803
QY 922 SAGCNNFSTQKIQYAT-----BFFEQSYLAIPASS-----WVDDFDIML----- 961
DB 804 -----YIVTQADYPNIQHLLDYLRHSFNVKYMLENKQLPKMWLHVFRDWLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPKDKFCPTVNSINCLKNCMSITMGS 1000
DB 858 AFDSDMWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVADAGI 913
QY 1001 VAPSVQEQHXYL-PMFLND-----RNNCKPKGGLAAY--STSVNLTSDGQ 1043
DB 914 INPSA--FYIYLTAVWSNDPVAYAAASQANIRPHRFEWHDK---ADYMPETRLRIPAAEP 968
QY 1044 VLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKVFGTDPAPFVFPYTTNVFY 1102
DB 969 IEYAQFPFVYNGLRDTSDFVBAIEKVRTICSNYTSGLSSYPNG-----YFP---LFW 1018
QY 1103 EOYLTILPEGLPMLSLCLVPTFPAVSCLLGLDLRSGLNLLSIVMLVDTVTGFMALWDIS 1162
DB 1019 EOYILGRHMLLFISVLAICTFLVCAVFLNPWTAGII--VMVALMTVELFGMGLIGIK 1077
QY 1163 YNAVSLINLVAAGVSGVFPVSHITRSP--AISTKPTWLERAKEATISMGSAVPAVAMTN 1220
DB 1078 LSAVPVILIASVGVGEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
QY 1221 LPIGLVLGLAKAQLQIIPFFRLNLLITLLGLHGLVFLPVLISYVG--PDVNPALALEQ- 1277
DB 1134 LLGLVLMAGSBDFIVRYFFAVLAITLITGLVNLGLVLLPVLISFFGYPYEVSPANGLNRL 1193
QY 1278 --KRABEAVAAMVWASCNHPHSRVSTADNIYVNHSPFEGSIKG 1317
DB 1194 FTPSPPEPPSVVRFAAMPFGHTH--SGSDSSDSEYSSQTTVSG 1233

; PRIOR APPLICATION NUMBER: US 60/077,723
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-140-5

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGAVGSGTQPLNEGVARCNESGDDVATCSCODCAASCPAARPOALDSTFYLGOMPGS 284
DB 24 PGRPAGGRRRTTGLRAAADPRD-----YLRHPSYCDAAFALEQI--- 65
QY 285 LVLIILCSFVAVVITLLVGFVRAPADKSKWDPK-----KGTSLDKLSFSTHTLL 337
DB 66 -----SKGKATGRKAPLWLRRAKFORLLFKLGCYIQKNC 98
QY 338 GQFFQCGTWTWASWPLTILVSVIPVVALAAGLVFTLTDPVELMSAPNSQARSKAFH 397
DB 99 GKF-----LVGLLIFGAFVAGLKAANLTNVEELMVEVGGVRSRELNT 143
QY 398 DQHFQGFPRFTN-QVILTAPNRSSRYDSLLGPKNPSGILDLLELE--LQRLRLHL 454
DB 144 RQKIGBEAMFNQMLTQTPKEEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QWSPAEQARNISLQDICYAP-----LNDPNTSLDYCCINSLLQYQNNRTLLTLTA 505
DB 191 YMYN----ROWKLEHLCKYSGBELITETGYMDQIIEYLPCLITPLDCFWEGAKLQSGTA 246
QY 506 NOTLMGO-----TSQVDWKHFLY-----CAN----- 527
DB 247 ---YLLGKPLRTWTFDPLEBELKKINTQVDSWEEMLNKAEVGHGMDRCPCLNPADPDC 304
QY 528 -----APL-----TFKQGT-----ALALSMA 544
DB 305 PATAPNKNSTKPLDMALVNLGCHGSLRKYMHQWBEELIVGGTVKSTGKLSAHALQTMF 364
QY 545 DYGAVPFPPLATGGYKGYSEABALIMTFSLNTPAGDPRLAQAKLWEEAFLEWRAPQ 604
DB 365 QLMTPKQMT---EHFKGYEY-----VSHINNE-----DKAAAILLEAWORTYVEVVHQS 411
QY 605 RMWAGMFQVTEAERSLEDEINRTTAEDLPFATSIVIVFLYISALGSYSSWSRVWDS 664
DB 412 AQNSTQKVLSTFT--TLLDDILKSFSDSVIRVASGYLLMLAYACLTM-----LRWD--CSKS 465
QY 665 KATGLGGVAVVLGAVMAAMGFFSYLGIRSSILVILQVVPFLVLSVGAADNIFVLEYQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLALGVGDDVFLAHAFSET 525
QY 725 PRPGEPREVHTGRALGRVAPNMLCSSEACFFLGALTPMPAVRTFALTSLGLAVILDF 784
DB 526 GQNKRIPEFDRTEGCKRTGASVALTSISNVTAFMAALIPALRAFSLQAQVAVVFN 585
QY 785 LLQMSAFVALLSLDKRQEAIBLVCCC-----VKQEL-----PPPG 822
DB 586 AMVLLIFPAILSMIDLYRREDRLDFCCFTSPCSRVIQVEQAVYTDTHDNTRYGPPPPY 645
QY 823 QCEGL----- 827
DB 646 SSHSAHETQITMQSTVQLRTEYDPTHVYTTAEPRSEISVQPTVTTQDTLSCQSPST 705
QY 828 -----LLGFFQKAYAPFLHWHITRGVLLLFALFGVSL 861
DB 706 SSTRDLLSQFSDSLHCLPEPPCTKTWLTSSFAEKHVPFLPKAKVWVIFLFLGLGVSL 765
QY 862 YSMCHISVGLDDELALPKDSYLLDFPLNRYFFVGAPVYFTTLYGNFSSAGNAICS 921
DB 766 YGTTVRDGLDLTDIVPRETREYDFIAAQKRYFSF-----YNN----- 803

QY 922 SAGCNNFSTQKIQYAT-----EPEOSYLAIPASS-----WVDDFIDWL----- 961
DB 804 -----YIVTQADYENIQHLLYDLHRSFSNVKYVWLEENKQLPKWMLHYFRDMLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDKFCPSTVNSLNCNCKMSITMGS 1000
DB 858 AFDSDWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQUT--KQRLVDADGI 913
QY 1001 VRPSVEQPHYL-PWFPLND-----RPNIKCPKGGLAAY--STSVNLTSDGQ 1043
DB 914 INPSA--FYIYLTAVNSNDPVAAYASQANIRHREWHDK--ADYMPETRURIPAAEP 968
QY 1044 VLASRFMAHYHFKPKNSQDYTEALRAARELANITA-DLRKVPCTDPAPFVFPVYITNVFY 1102
DB 969 IEYAGFPFYLNGLRDTSDFEAEIKVKTICSNVTSGLSSYPNG-----YFP---LFW 1018
QY 1103 EOYLITLPEGLPMLSCLVPTFVAVSCLLGLDLRLSGLNLLSIVMLLVDTVGVPMALWDIS 1162
DB 1019 EOYIGLRHMLLFISVLACTIONLCAVFLNPTAGII-VMVLAALMTVELFGMGLIGIK 1077
QY 1163 YNAVSLINLVSAGMSVRFVSHITRSP--AISTKPTWLERAKEATISMGSAVPAGVAMTN 1220
DB 1078 LSAVPVILIASVGIGVEFTVHALAFIATGDKN---RRVLALEHMFAPVLDG-AVST 1133
QY 1221 LFGILVLGLAKAQLIQIFFRNLMLLITLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277
DB 1134 LAGVLMLAGSBDFIVRYFFAVLAITLITGLVNLGLVLLPVLSPFGPYEVSFANGLANRL 1193
QY 1278 --KRAEAVAAMVASCNHPHSRVSTADNIYVNHSPGSIKG 1317
DB 1194 PTPSPPEPPSVVRFPAMPCHTH--SGSDSSDSEYSSQTTVSG 1233

RESULT 10

US-08-918-658-19
; Sequence 19, Application US/08918658
; Patent No. 6429354
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/918,658
; FILING DATE: 22-Aug-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/656,055
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/540,406
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:	
LENGTH: 1447 amino acids	
TYPE: amino acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: protein	
SEQUENCE DESCRIPTION: SEQ ID NO: 19:	
US-08-918-658-19	
Query Match	9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity	22.0%; Pred. No. 1.1e-51;
Matches 300; Conserved 181; Mismatches 460; Indels 421; Gaps 47;	
QY	225 PQAVGSGIQLINEGVARNESQGDVATCSCDCAACPAIAIPAOLDSTFYLGQMPGS 284
DB	24 PGRAGGRRRTTGLRRAAPDRD-----YLRPSYCDAAFALEQI--- 65
QY	285 LVLIILCSFVAVTILLVGRFVARDKSKWDPK-----KGTSLSDKLSFSTHLL 337
DB	66 -----SKGATGKAKPLMRAPQRLFLKLCGYIQKC 98
QY	338 GQFFQGWGWASWPLTILVLSVIPWALAAGLVTELTDPVELWSAPNSQARSEKAFH 397
DB	99 GAF-----LVGLLIFGAFAGVGLKAANLETVNELWVEVGGVRSRELYNT 143
QY	398 DQHFQFPFRTN-QVILTAPNRSRYDSLLGPKNPSGILDLDLLELLE--LQERLRL 454
DB	144 RQKIGSEAMFNQMIQTKBEG-----ANVLTTAQLQHLDSALQASRVHV 190
QY	455 QWSPQAQRNISLQDICYAP-----LNPDTSLYDCCINSLLQVFNRTLLTLTA 505
DB	191 YMYN-----RQWKLEHLCYKSGELITETGYMDQIIEYLPCLITPLDCFWGAKLQSGTA 246
QY	506 NOTLMGQ-----TSQVMDKDFLY-----CAN----- 527
DB	247 ---YLLGKPLRWTNPDPLFLELKLKINYQVDSWEMLNKAEGVGHGYMDRCLNPADPDC 304
QY	528 -----APL-----TFKQGT-----ALALSQWA 544
DB	305 PATAPNKNSTKPLDMALVNGGCHLSRKYMHQWELIVGGTVKNSTGKLVSAHALQTMF 364
QY	545 DYCAPVFPFLAIGYKGDYSBAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEENRAFO 604
DB	365 QLMTPKQMY---EHFGYEF-----VSHINNE-----DKAAALEAWQRTYVEVHQSV 411
QY	605 RRMAGMFQVFTTAERSLEDEINRTAEDLPIFATSYIVIFLYISLALGYSNSRVMVDS 664
DB	412 AQNSTQKVLSTFT--TTLDLILKSFSDSVIRVASGYLLMLVACLTM---LRWD--CSKS 465
QY	665 KATLGLGGVAVLVGAVNAAMGFYSYLGIRSSLVILQVVPFLVSLVGCADNIFIFVLEYQRL 724
DB	466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFALGVGVDDVFLLAHAFSET 525
QY	725 PRPGEPREHIGRALGRVAPSMLLCSLSAICFFLCAITPMPAVRTFALTSLAVILDF 784
DB	526 GQNKRIPEDETCRLKTCASVALTSGISNWTAFMAALIPALRAFSLQAQVVVVF 585
QY	785 LLOMSAFVALLSLDSKQEARSLDVCC-----VKPEL-----PPPG 822
DB	586 AMVLLIFAILSLMDLYRREDRLDIFCCFTSPCVSRVQVPEQAYTDTHTNTRYSPPPY 645
QY	823 QGEGE----- 827
DB	646 SSHSFAHETITWQSTVQLRTEYDPHTVHTVYTAEPSEISGVOPVTVTQTLSCQSPST 705
QY	828 -----LLGPFQKAYAPFLHWTIRGVVLLFLFALFGVSL 861
DB	706 SSTFDLLSQPSSLSHCLPECTKWTLSSFAEKHYAPFLKPKAKVWVIFLGLLGVSL 765
QY	862 YSMCHISVGLDQELALPROSVLLLDYFLFLNRYFVGAPVFTVTLGNFSFSEAGMAICS 921
DB	766 YGTTVRDGLDLTDIVPRETREYDFIAAQKYSF-----YNN----- 803
RESULT 11	
US-09-724-631-19	
; Sequence 19, Application US/09724631	
; Patent No. 6551782	
GENERAL INFORMATION:	
APPLICANT: SCOTT, MATTHEW P	
GOODRICH, LISA V	
JOHNSON, RONALD L	
TITLE OF INVENTION: Patched Genes and their Use	
NUMBER OF SEQUENCES: 19	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert	
STREET: Four Embarcadero Center, Suite 3400	
CITY: San Francisco	
STATE: CA	
COUNTRY: US	
ZIP: 94111	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Floppy disk	
COMPUTER: IBM PC compatible	
OPERATING SYSTEM: PC-DOS/MS-DOS	
SOFTWARE: Patent In Release #1.0, Version #1.30	
CURRENT APPLICATION NUMBER:	
FILING DATE: 28-Nov-6551782-2000	
CLASSIFICATION: <Unknown>	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: US/08/656,055	
FILING DATE: 1996-05-31	
APPLICATION NUMBER: 08/540,406	
FILING DATE: <Unknown>	
ATTORNEY/AGENT INFORMATION:	
NAME: Rowland, Bertram I	
REGISTRATION NUMBER: 20015	
REFERENCE/DOCKET NUMBER: a60190-1	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: 415-781-1989	
TELEFAX: 415-398-3249	
INFORMATION FOR SEQ ID NO: 19:	
SEQUENCE CHARACTERISTICS:	

```
; ; LENGTH: 1447 amino acids
; ; TYPE: amino acid
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: protein
; ; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-724-631-19

Query Match          9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PCGAVSGIGLOPLNEGVARCNESGDDVATCSODCAASCPAIPARQALDSTFVYLGMPGS 284
DB 24 PCRAGGRRRTTGGRLRAAARD-----YLRPSYCDAAFALEQI--- 65
QY 285 LVLIILILCSFAVAVTILLVGRFVAPARDKSKWDPK-----KGTSLSDKLSPSTHTLL 337
DB 66 -----SKGKATGRKAPLWLRAKFORLLFLKGCYIQKNC 98
QY 338 GQFFQCGWTWASWPLTILVSVIPVVALAAGLVFTTELTPVVELMSAPNSQARSEKAPH 397
DB 99 GKF-----LVUGLLIFGAFVGLKAANLETVNEELWVEVGGVSRSLNTY 143
QY 398 DQHFQPPFTN-QVILTPANRSYRYSLLGPKNPSGILDLDLLELE--LQERLRL 454
DB 144 ROKIGSEAMFNPLMTQTPKEG-----ANVLTTEALIQHLDLSALQASRVHV 190
QY 455 QVMSPEAQNISIDICYP-----LNPNTSLYDCCINSLLIOYQNNRTLLTLTA 505
DB 191 YMIN-----KQWLEHLCYKSGELITETGYMDQIIEYLYPCLITPLDCFWEGAKLQSGTA 246
QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
DB 247 --YLLGKPLRWTNFDPLBELKKNYQVDSWEEMLNKAEVGHGYMDRCPNLPADPPC 304
QY 528 -----APL-----TFKQGT-----ALALSCWA 544
DB 305 PATAPNKNSTKPLDMALVNGGCHGLSRKMYHQBELIVGGTVKNSTGKLSAHALQTMF 364
QY 545 DYGAVPFPPLAIGGYKDYSEALIMTFSLNNYPAGDPRLAQAKLWEAEFLEMRAQ 604
DB 365 QLMTPKQMT---EHFKGYEY-----VSHINNE-----DKAAALAEAWQRTTVEVHQS 411
QY 605 RMAGMFOVFTAERSLDEINRTAEDLPIFATSYIVIFLYISALGSYSSWSRWVDS 664
DB 412 AQNSTQKVLSTFT--TFLDILKSFSDSVIRVASGYLLMAYACLTM---LRWD--CSKS 465
QY 665 KATLGLGVAVVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVGADNIFIVLEYQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGDDVFLAHAFST 525
QY 725 PRPGEPRVHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTFALTSGLAVIDP 784
DB 526 GQNKRIFFEDRTGECIKRTGASVALTSISNTVTAFFMAALIPALRAFSLQAAVVVNF 585
QY 785 LIQMSAFVALLSLDKRQBSRLDVCCC-----VKQEL-----PPPG 822
DB 586 AMVLLIFPALLSNDLYRDRDLDFCCFTSPCVSRVIOVEFQAYTDHNDTRYSPPPY 645
QY 823 QGEG----- 827
DB 646 SSSPAHETQITMQSTVQLRTEYDPHTHVYVYTAEPBSISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGFFKAYAPFLHWTITRGVLLLFLALFGVSL 861
DB 706 SSTRDLSQFSSSLHCLBPCTKWTLSFAEKHAPFLPKAKVWVIFLFLGLGVSL 765
QY 862 YSMCHISVGLDQELPKDSYLLDYFLNRYFVGAPVYFVTLTGYNFSSRAGNAICS 921
DB 766 YGTRVRDGLDLDTDIVPRETRYDFIAQFKFSP-----YMN----- 803
QY 922 SAGCNNFSPTQIOYAT-----EPPEQSYLAIPASS-----WVDDFDWL----- 961
```

```
DB 804 -----YIVTQADYPNIQHLLYDLHRSFSPNVKYVMLEENKQLPKMWLHYFRDLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKKDFCPSVTNSLNCNCKMSITMGS 1000
DB 858 AFDSWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQUT--KORLVDADGI 913
QY 1001 VRPSVEQFHXYL--PWFLND-----RPNIKCPKGLGAAY--STSVNLTSDGQ 1043
DB 914 INPSA--FYIYLTAWVNDPVAAYASQANIRPHRDEWHDK---ADYMPETRLRIPAAEP 968
QY 1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA-DLRKVPCTDPAFVFPYTIINVFY 1102
DB 969 IEYAFPPFYINGLRDTSDFEAEIKVRTICSNYTSGLSSYPNG-----YDF--LFW 1018
QY 1103 EGYLTILPEGLPMLSLCLVPTFAVSCLLGLDLRSLGLNLLSVMLLVDTVGFMAWDIS 1162
DB 1019 EGYIGLRHWLLLFISVVLACTFLVCAVFLNPTAGII--VMVLAJMTVSLFGMGLIGIK 1077
QY 1163 YNAVSLINLIVSAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAVAGVAMTN 1220
DB 1078 LSAVPVILLIASVGIGVETVHVVALAFLTAIGDKN---RRAVLALHEHMFAPVLDG-AVST 1133
QY 1221 LRGIIVLGLAKAQLIQIPFRNLNLLITLGLLHGLVFLPVILSYVG--PDVNPALALEQ- 1277
DB 1134 LLAGVLMLAGSEPDFIVRYFFAVLAILITLGLVGLVLLPVLLSFFGYPYEVSPANGLNRL 1193
QY 1278 --KRAEEAVALVMAVSCNHPSPVSTADNIYVNHSPGSGIKG 1317
DB 1194 PTPSPPEPPSVVRFAWPPGHTH--SGSDSSDSEYSSQTTVSG 1233

RESULT 12
US-08-954-701A-19
; Sequence 19, Application US/08954701A
; Patent No. 6610507
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-701A-19
```

QY	1001	VRPSVEQPHKYL-PWFLND-----RPNIKCPKGGGLAAY--STSVNLTSDDGQ	104
Db	914	INPSA--FYIYLTAWSNDFVAVAAASQANIRPHRPWWHDK---ADYMPETRLRIIPAAEP	968
QY	1044	VLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKVPGTDPAFEPVPTITNVFY	1102
Db	969	TEYAQFPYLNGLKXDTSDFEAIEKVRTICSNYTSIGLSYPNG-----YFF-----LFW	1018
QY	1103	EQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRLSGLNLNLLSVIMILVDTVGFMALWDIS	1162
Db	1019	EQYIGLRHLLLFISVVLACTFLCAVFLNLPWTAGII-VNVLALMTVELFGMGLIGIK	1077
QY	1163	YNAVSLINLVAAGMSVEFVSHITRSP--AISTPTWLERAKEATISWGSNAVPAGVAMTN	1220
Db	1078	LSAIVPVLIIASVGIQVEFTVHVALAFLTAIGDKN---RRAVLAHEHMFAPVLGD-AVST	1133
QY	1221	LRGILVLGLAKAQLIQIIFPRLLNLLITLLGLLHGLVFLPVLISYVG--PDVNPALALEQ-	1277
Db	1134	LUGVLMLAGSDFIVRYFFAVLAILITLGVNLGVLLVLLSFFGYPVPSPANGLNRL	1193
QY	1278	--KRAEEAVAAVMVASCNPHPSRVSTADNIYVNHSPFEGSING	1317
Db	1194	PTPSPPEPPSVVRFAMPGGHTH--SGSDSDSEYSSOTTVSG	1233
RESULT 13			
US-09-807-007-6			
; Sequence 6, Application US/09807007			
; Patent No. 6881833			
; GENERAL INFORMATION:			
; APPLICANT: ZAPHIROPOULOS, Peter et al.			
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY			
; FILE REFERENCE: 2921-0130P			
; CURRENT APPLICATION NUMBER: US/09/807,007			
; CURRENT FILING DATE: 2001-04-06			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 6			
; TYPE: PRT			
; LENGTH: 1447			
; ORGANISM: Homo sapiens			
US-09-807-007-6			
Query Match 9.4%; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47			
QY	225	PGQAVSGIQPLNEGVARCNESQGDVATCSCQDCAASCPAIPARQALDSTFYLGQMPGS	284
Db	24	PCRAGGRRRTTGLRRAAAPDRD-----YLRPSYCDAAFALEQI---	65
QY	285	LVLIIILCSVFAVVITLLVGFVRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL	337
Db	66	-----SKGKATGRKAPLWLRKAFQRLLFKLCGYIQKNC	98
QY	338	GOFFQCGMTWASWPLTILVLSVIPVVALAAGLVFTELTDPVELWSPAPNSQARSEKAPH	397
Db	99	GKF-----LVVGLLIIFGAPAVGLKAAANLETNVEELWVEVGGVRSRELNYT	143
QY	398	DQHFGPPFRTN-QVILTAENRSSRYDSLLGLPKNFSGILDLDLLELE--LQERLRLH	454
Db	144	RQKIGEEAMFNPMIQTPEEG-----ANVLITTEALLQHLDSALQASRVHV	190
QY	455	QVWSPEAQRNISLQDICYAP-----LNPDNTSLYDCCINSILQYFQNNRTLLLLTA	505
Db	191	YMYN----RQWLEHLCHYKSGELITETGYMDQIIIEVLPCLITPLDCEWEGAKLQSGTA	246
QY	506	NOTLMGO-----TSQVDWKDHFY-----CAN-----	527
Db	247	--YLLGKPLRWTFNDFLEBLELKKINYQVDSWEEMLNKAIEVGHGYMDRPCLNADPDC	304
QY	528	-----APL-----TFKQGT-----ALALSCMA	544
Db	305	PATAPNKNSTKPLDMALVNLGGCHLSRKYMHQBELIVGGTVKNSTGKLSAHLQTMF	364
QY	545	DYGAIPPELAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEBMAFQ	604
Db	365	QLMTPKQMI---EHPKGYEY-----VSHINWNE---DKAAAILLEAMQRTIVVVHQS	411
QY	605	RRMAGMPQVTTAERSLEDEINTTAEDLPFATSYVIVFLYISALGYSWSRWVDS	664
Db	412	AQNSTQKVLSTFT--TTLDDILKFSQSVIRVASGYLLMLAYACLTM---LRWD--CSKS	465
QY	665	KATLGLGGVAVLGAWMAAGFFSYLGIRSSLVILQVFPFLVLSVGCADNIFIFVLEYQRL	724
Db	466	QGAAGVAGVLLVALSVAAGLGLCSLGISFNATTTQVLFPLALGVGVDDVFLLAHAFSE	525
QY	725	PRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFLFGALTPMPAVRTFALTSLGLAVILDF	784
Db	526	GQNKRIIPEDRTGECRLKRTCASVALTISINVTAFMAALIPALRAFSLQAAVVVWFNF	585
QY	785	LQWSAFVALLSDSKQBSRLDVCC-----VKQEL-----PPRG	822
Db	586	AMVLLIIFAILMDLYRREDRLDIFCCFTSPCVSRVIVQEPQAYTDDHNTRYSPPPPY	645
QY	823	QGEGL-----	827
Db	646	SSHSEFAHETQITMSTVQLRTEYDPHTVYVYTTAEPRSEISVQPVTVTQDTLSCQSPE	705
QY	828	-----LLGFFQKAYAPFLHLWITRQVWVLLFLALFGVSL	861
Db	706	SSTRDLSQFSDSLHCLPECTKTWTLSSFAEKHYAPFLKPKAKVWVIFLFLGLLGVSL	765
QY	862	YSMCHISVGLDOELALPKDSYLLDYFLFLNRYEFGVAPVYFVTTLGVNFSSEAGMNAICS	921
Db	766	YGTTRVRDGLDITVLPRETRYDFIAAQFYKFSF-----YNN-----	803
QY	922	SAGCMNFSFTQKIQYAT-----EPFEQSYLAIPASS-----WVDDFDIWL-----	961
Db	804	-----YIVTKADYPNIOHLLYDLHRSFSNVKVMLEENKQLPKWMLHYFRDWLQLOD	857
QY	962	-----TPSS-----CCRELYISGNKDKFCPSVTNSINCLNCWMSITMGS	1000
Db	858	AFDSDWETGKIMPNYKNGSDGVLYAKLLVQTGSRDK---PIDISQUT--KQRLVADG	913

Db 305 PATAPNKNSTKPLDMALVNLGCHGLSRKYMHWQBELIVGGTVKNTGKLVSAHALQTMF 364
QY 545 DYGAVPFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLMEAEFLMEARAFQ 604
Db 365 QLMTPKQMY---EHPKGYEY-----VSHINWNE-----DKAAAILLEAWQRTVVVHQSV 411
QY 605 RMWAGHPQVTPAERSLDEINRTAEDLPFATSYIVIFLYISALGYSYSSWRVWDS 664
Db 412 AQNSTQKVLSEFTT-TTLDDILKSPSDSVIRVASGYLLMLAYACLTM---LWMD--CSKS 465
QY 665 KATGLGVAVVGLVGAAMAGFFSYLGRSSVILQVVPFLVLSVGCADNIFIVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATTVQLPFLALGVGVDDVFLAHAFSET 525
QY 725 PRPGEPREHVHTRALGRVAPSMGLCSLSEAI CFFLGALTMPAVRTFALTSLGLAVILDP 784
Db 526 GQMKRIPPEDRTGECLKRTGASVALTSGISNVTAFMAALIPALRAFSLQAQVVVFN 585
QY 785 LLOMSAFVALLSLDKRQASRLDVCCC-----VKQEL-----PPPG 822
Db 586 AMVLLIFPAILSMDLRYRREDRLDFCCFTSPCSRVIQVEPQAYTDDHNTRYSPPPY 645
QY 823 QCEGL----- 827
Db 646 SSHSAHETQITMQSTVQLRTEYDEPHVYVYTTAPRSEISVQPVTVTQDTLSCOSPST 705
QY 828 -----LLGFFOKAYAPFLHWHITRGVLLLFALFGVSL 861
Db 706 SSTDRLLSQFSSSLHLEPPCTKWTLSSFAEKHYAPFLPKAKVWVIFLGLLGVSL 765
QY 862 YSMCHISVLGDELALPKDSYLLDFLNFIRVEGAPVYVTTTLYGNPSSAGNMAICS 921
Db 766 YGTRVRDGLDITDVPRETRYDETAQPKYFSF-----YNN----- 803
QY 922 SAGCNFSPTQIQYAT-----EPPEOSYLAIPASS-----WVDDRIDML----- 961
Db 804 -----YIVTQADTPNQLHYDLHRSFNVKYLWLENKOLPKMWHLYFRDMLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDFPCPSTVNSLNCNLCMSITMGS 1000
Db 858 AFSDSWETGKIMPNNYKNGSDGGLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
QY 1001 VRPSVEQFKYL--PWFLLND-----RNKICPKGGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYIYLTAWNSDPVAYAAQANIRPREWVHDK---ADYMPETRLRIPAAEP 968
QY 1044 VLASRFMAVHKPLKNSQDYTEALRAARELAANITA-DLRKVPCTDPAVEPEYTTITNVFY 1102
Db 969 IEYAQFPFYLNGLRDTSDFVEAIEKVRTICSNYTSGLSSYPNG-----YPF-----LFW 1018
QY 1103 EQYLTLPEGLPMLSLCIVPTFAVSCLLIGLDRGLNLNLSIVMLVDTVGFMAWDIS 1162
Db 1019 EQYIGLRHWWLLFISVVLACTFLVCANVFLNPWTAGII-VWVLAJMTVELFCMGLIGIK 1077
QY 1163 YNAVSLINLVSAVMSVRFVSHITSP--AISTKPTWLERAKBATISGMSAVFAGVAMTN 1220
Db 1078 LSAPVPWILLASGVIGVEFTVHALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
QY 1221 LPGILVLGAKAQLQIOPFRNLNLLITLLGLHGLVFLPVILSYVG--PDVNPALALSO- 1277
Db 1134 LLGLVLMAGSEPDFVIRYFAVLAITLITGLVGLVLLVPLLSFFGYPPEVSPANGLNRL 1193
QY 1278 --KRAEAAVAAVWASCPNHPRSVSTADNIYVNHSPFGSIKG 1317
Db 1194 PTPSPPEPPSVVRPAMPFGHTH--SGSDSSDSEYSSQTTVSG 1233

RESULT 14

US-09-754-032-19
; Sequence 19, Application US/09754032
; Patent No. 6921646
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P

GOODRICH, LISA V
JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
CITY: Four Embarcadero Center, Suite 3400
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/754,032
FILING DATE: 03-Jan-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/540,406
FILING DATE: 06-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-754-032-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;

Best Local Similarity 22.0%; Pred. No. 1.1e-51;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQAVGSGIQPLNEGVARCNESQGDVATCSQDCAACPAARQALDSTFVLGMPGS 284

Db 24 PGRPAGGRRRTTGLRRAAADRD-----YLRPSYCDAAAFALQI--- 65

QY 285 LVLIILCSVFAVVTILLVGFVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337

Db 66 -----SKGATGRKAPLMLRAKQRLFLKLGCIQKNC 98

QY 338 GQPFQGWGTWASWPLTTLVLVSVIPVALAAGLVTELTTPDVELWSPNSQARSEKAPH 397

Db 99 GKF-----LVVGLLIFGAFVGLKAAANLETNVVEELWVEVGVRVSELNVT 143

QY 398 DOHGFEPFRTN-QVLTAPNESSRYSDLSLLGPKNFSGILDLLALLELE--LQERLRHL 454

Db 144 RQKIGEEAMFQPMIQPKKEG-----ANVLTTEALQHLDSALQASRVHV 190

QY 455 QVMSPEAQRNLSLODICYAP-----LNPNDTSLYDCCINSLLQYFQNNRTLLLLTA 505

Db 191 YWYN---ROWKLEHLCYKSGELITETGYMDQIIEYLYPCLIIITPLDCFWGAKLQSGTA 246

QY 506 NQTLMGQ-----TSQVMDKHFLY-----CAN----- 527

Db 247 --YLLGKPELRWTNPDPLFLEELKKNYQVDSWEMLNKAEGVGHYMDRPLCLNPAEDFC 304

QY 528 -----APL-----TFKDG-----ALALSCWA 544

Db 305 PATAPNKNSTKPLDMALVNLGCHGLSRKYMHWQBELIVGGTVKNTGKLVSAHALQTMF 364

QY 545 DYGAVPFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLMEAEFLMEARAFQ 604

Db	365	QLMTPKQMY---BHFKEGYEY-----VSHINWNE-----DKAAAILAEWQRTVEVVHQSV 411
Qy	605	RRWAGHFQVPTTAERSLEDEINRTTAEDDIPFATSVIVIFLYISLALGSSYSSKSRVWVDS 664
Db	412	AQNSTOKVLSFTT--TTLDDTLKFSQSVSVIRVASGYLLMLAYACLTM-----LRWD--CSKS 465
Qy	665	KATLGLGGVAVLGAWMAAGPFYSILGRSSVILQVWPELVLSVCADNIFIVLVEYQRL 724
Db	466	QGAVGLAGVLLVALSVAAGLUGLCSLIGISNAATTQVLPFLAGVGDDVFLLAHAFSET 525
Qy	725	PRRPGPREVHIGRALGRVAPSMLLCSLSSEAI CFFLGCALTTPMPAVRTALTSLGAILDP 784
Db	526	GQNKRPEDRTGECKRTGASVALTISINVTAFMAALIPFALRAFSLQAADVVFNF 585
Qy	785	LLOWSAFVALLSDSKRQASRLDVCCC-----VKPEL-----PPPG 822
Db	586	AMVLLIPPAILSMDLYRREDRLDIFCCFTSPCVSRVIOPEQAYTDTHNTRYSPPPY 645
Qy	823	QGEGL-----827
Db	646	SSHSPAHETQTMQSTVOLREYDPHTVHYTTABPRSEISQPVTVTQDTLSCQSPST 705
Qy	828	-----LLGFFQKAYAPFLLHMTGRGVLLLFLALFGVSL 861
Db	706	SSTRDLLSQSDSLHCLBPPCTKWTLSFPAEKHYAPFLLKPAKVVVVFLGILGVSL 765
Qy	862	YSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGPVYFVTTLGVNFSSEAGMAI CS 921
Db	766	YGTTRVDGLDLTDIVPRETREYDFIAAQPKYFSF-----YNM-----803
Qy	922	SAGCNFSFTKIOYAT-----EPPEQSYLAIPASS-----WVDDFIOWL-----961
Db	804	-----YIVTQADYPIQHLLYDLHRSFNVKYVMLEENKQLPKWMLHYFRDWMLOGLQD 857
Qy	962	-----TPSS-----CCRLYISGPNKDKFCPSVTNSLNCNKCMSITMGS 1000
Db	858	AFSDSWFTGKMPNPKYKNGSDGVLAYKLLVQTGRDK--PIDISQUT--KQRLVDADGI 913
Qy	1001	VRPSVEQPHKYL--PWFLND-----RPNIKCKPKGGILAA--SSTVNLTSDDG 1043
Db	914	INPSA--FYIYLTAVVSNDPVAYAASQANIRPHRPWVHDK--ADYMPETRLRIPAEEP 968
Qy	1044	VLASRFWAYHKPLKNQODYTEALRAARELANITA--DLRKVGPDPFAFEPVPTITNVFY 1102
Db	969	IEYAQFPFYLNGRLDTSFDEAIEKVRTICSNYTSLSGLSSYPNG-----YPF-----LFW 1018
Qy	1103	BOYLTILPEGLFMLSCLVPFPFAYSCLLGLDLRSGLLNLLSIVMILAVDIVGMAIWDIS 1162
Db	1019	BOYGLRHWLLLFISVVLACTFLVCVFLNLPWPTAGII--VMVLAIMTVELFGMMGLIGIK 1077
Qy	1163	YNAYSLINLVAAGVMSVEFVSHIRSF--AISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db	1078	LSAVPVLITASVIGVEFTVHVVALAFLTAGDN-----RRAVLALEHMFAPVLDG--AVST 1133
Qy	1221	LPGLIIVLGLAKAQLTIQTPFFRLNLLITLLGLLHGLVFLPVILSYVG--PDVNPALALEQ-- 1277
Db	1134	LLGVMLAGSEFDFIVRYFFAVAILITLIGVLNGLVLLPVLLSFFGYPVEVSPANGLNRL 1193
Qy	1278	--KQABEAAVAAVMVASCPNHPSPRSTADNIYVNHISFEKSGK 1317
Db	1194	PTSPPEPPSVVRFAMPBGHTH--SGSDSDSEYSSOTTVSG 1233

```

; TITLE OF INVENTION: PATCHED GENES AND USES RELATED THERETO
;
; FILE REFERENCE: CIBT-P04-203
; CURRENT APPLICATION NUMBER: US/08/916,140
; CURRENT FILING DATE: 1997-08-21
;
; PRIOR APPLICATION NUMBER: US 08/656,055
; PRIOR FILING DATE: 1996-05-31
;
; PRIOR APPLICATION NUMBER: US 08/540,406
; PRIOR FILING DATE: 1995-10-06
;
; PRIOR APPLICATION NUMBER: US 08/319,745
; PRIOR FILING DATE: 1994-10-07
;
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 19
;
; LENGTH: 1447
;
; TYPE: PrT
; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Primer
;
; US-08-916-140-19

```

[illegible]


```
Db 646 SSHSFAHETQITMQSTVQLRTEYDHPHTHYTAAERSEISVQPVTVTQDTLSCQSPEST 705
QY 828 -----LLGFFOKAYAPFLHMITRGTGVLILLFLALFGVSL 861
Db 706 SSTRDLLSQSDSSHLCLPCTKWTLSFPAEKHYAPFLPKPAKVWVFLFGLLGVSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEGAPVYVFTTGLGNFSSSEAGMNAICS 921
Db 766 YGTRVRDGLDLDIVPRETREYDFAAQKPYFSF-----YNN----- 803
QY 922 SAGCNFSFTQKIQYAT-----ERPEQSYLAIPASS-----WVDDFDIML----- 961
Db 804 -----YIVTQADYNIQHLLYDLHRSFNSVYVMLEENKQLPKWMLHYFRLWGLQLD 857
QY 962 -----TPSS-----CRLYISGNKDKFCSTVNSLNCIKNCMSITMGS 1000
Db 858 APDSDWETGKIMNNYKNGSDGVLAYKLLVQTSRDK--PIDISQLT--KQLVDADGI 913
QY 1001 VRPSVQFHKYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLTSDDG 1043
Db 914 INPSA--FYIYTAWSNDPVAAYASQANIRPHRPWVHDK--ADYMPETRLRIPAAER 968
QY 1044 VLASRWAYHKLKNSQDYTEALRAARELANITA--DLRKVPCTDPAPEPPTIINVFY 1102
Db 969 IEYAQPFYGLNGLRDTSDFVEALEKVRTICSNYSIGLSYPNG-----YPP-----LFW 1018
QY 1103 EQYLTLPBGLFMSLCLVPTFAVSCLLGLDLRLSGLNLLSIVMLIVDTVGPMLWDIS 1162
Db 1019 EQYIGLRHMLLPIFISVVLACTFLVCVFLNPNWTAGII--VWVLALMTVELFGMGLIGIK 1077
QY 1163 YNAVSLINIVSAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAVPAGVAMTN 1220
Db 1078 LSAVPVILLIASVIGVEFTVHVALAFLTAIGDKN--RRAVLALRHMPAPVLDG-AVST 1133
QY 1221 LPGILVGLAKAQLIQIFRRLMLLITLGLLHGLVPLVILSYVG--PDVNPALALEQ- 1277
Db 1134 LGLVLMAGSEFDIVRYFPAVALAITLGLVNLGLVLLPVLFSFFGYPFVSPANGLNRL 1193
QY 1278 --KRAEEAVALVMAVSCPNHPSRVSTADNIYVNHSPGSIKG 1317
Db 1194 PTPSPPPSVVRFPMPGHTH--SGSDSDSEISSQTTVSG 1233

RESULT 16
PCT-US95-13233-19
; Sequence 19, Application PC/TUS9513233
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13233
; FILING DATE: 06-OCT-1990
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
```

```
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-13233-19

Query Match 9.4%; Score 651.5; DB 4; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGQAVSGGIQPLNGBVARNESQGDVATCSQDCAASCPAIAAPQALDSTFVGLQMPGS 284
Db 24 PGRPAGGRRRTTGLRRAAADPRD-----YLRPSYCDNAFALEQI--- 65
QY 285 LVLIILCSVFAVVVITLLVGFVRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWLRKAFQRLFLKGLCYIQKNC 98
QY 338 GQFFQCGWTWASWPLTILVSVIPVVALAAGLVFTTELTDTPVELMSAPNSQARSKAPH 397
Db 99 GKF-----LVVGLLIFGAPAVGLKAANLETNVEELWVEVGGRVSRBLNYT 143
QY 398 DQHFPPFRTN--QVILTAPNRSSRYDSSLGPKNPSGILDLDLLELLE--LQERLRHL 454
Db 144 RKIGEEAMPNQLMIQTPEEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QVMSPEAQBNISLDIQCYP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 191 YVYN---RQWKLHLCYKSGELITETGYMDQIIBLYPCLITPTLDCFWEGAKLQSGTA 246
QY 506 NOTLWGO-----TSQVDKWKOHFLY-----CAN----- 527
Db 247 ---YLLGKPLRWNTFPDLFLEELKKINYQVDSWEMLNKAEGVGHGMDRCPCLNPADPDC 304
QY 528 -----APL-----TFKQGT-----ALALSQWA 544
Db 305 PATAPNKNSTKPLDMALVNLGGCHLSRKYMWQBELIVGGTVKNSTGKLVSAAHLOTMP 364
QY 545 DYGAFFVFPFLAIGYKGYDSEAEALIMTFSLNYPAGDPRLAQAQKWEAFLEENRAFQ 604
Db 365 QLMTPKQWY---EHFKGYEY-----VSHINWNE---DKAAALEAMQRTVEVVHQSV 411
QY 605 RRMAGHFQVTTAERSLEDEINRTTAEDLPITATSYIVIFLYISIALGYSYNSRWVVD 664
Db 412 AQNSTQKVLSPFT--TTLDDILKSPSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
QY 665 KATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVVPPLVLSVGADNIFIFVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLFALGVGVDDVFLAHAFST 525
QY 725 PRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTTALTSLGLAVILDF 784
Db 526 CQNKRIPEEDRTGECIKRTGASVALTISNTVTAFTMAALIPALRAFSLQAQVAVVNF 585
QY 785 LLQMSAFVALLSLDSKROEASRLDVCCC-----VKQDEL-----PPPG 822
Db 586 AMVLLIFPAILSMOLYRREDRLDIFCCFTSPCSRVIQVEPQAVTDTHDNTRYSPPPY 645
QY 823 QDEGL----- 827
Db 646 SSHSFAHETQITMQSTVQLRTEYDHPHTHYTAAERSEISVQPVTVTQDTLSCQSPEST 705
QY 828 -----LLGFFOKAYAPFLHMITRGTGVLILLFLALFGVSL 861
Db 706 SSTRDLLSQSDSSHLCLPCTKWTLSFPAEKHYAPFLPKPAKVWVFLFGLLGVSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEGAPVYVFTTGLGNFSSSEAGMNAICS 921
Db 766 YGTRVRDGLDLDIVPRETREYDFAAQKPYFSF-----YNN----- 803
```


Db 973 FVEATEKRVICNNYTSLSGSSYPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIYVILVDTVGFMAWLDISYNAVSLINLSAVGMSVE 1180
Db 1023 ACTFLVCAVFLNPTAGII-VNVLALMTVELFGWMLGIGIKLSAVPVVILLASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVFAVAVMTNLPGLILVLGLAKAQIQLF 1238
Db 1082 FTVHVAFALTAIGDKN---HRAMLALEHMFAPVLDG-AVSTLLGLVLMLAGSEFDFIVRY 1137
QY 1239 FFRNLNLTLLGLLHGLVFLPVLISYVG--PDVNPALALEQ---KRAEBAVAAMVASC 1293
Db 1138 FPAVALITLVGLVGLVLLPVLISFFGCPCEVSPANGLNRLPTSPSPPPSVVRPAVPP 1197
QY 1294 NHPSRVS-TADNIYVNHSPESI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISEBELROYEAQOGAG 1234

RESULT 18

US-08-656-055-10
; Sequence 10, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406

FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-10

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
QY 356 LVLSVTPVVALAAGLVFTLTTPVLSAPNSQARSEKAFDHFQPPFRTN-QVILTA 414
Db 88 LVVGLLIFGAFVGLKAANLENTVELWVEVGRVSRLENTYTRQKIGBEMFPQLMIQT 147
QY 415 PNRSSRYRYSLLGPKNFSGILDLLLELLE--LQERLRHLQVMSPEAQRNISLDICY 472

Db 148 PKBEG-----ANVLTTALLQHLDSALQASRVHVVMYN-----RQWLEHLCY 190
QY 473 AP-----LNPONTSLYDCCINSILQYFQNNRTLLALLTANQTLMO-----512
Db 191 KSGELITETGYMDQIIIEYLYPCLIIITPLDCFWEKAKLQSGTA--YLLGKPPPLRWTFDPL 248
QY 513 -----TSQVDKDHFLY-----CAN-----APL-----530
Db 249 BPLEBELKKINYQVDSWEEMLNKAEBVGHGYMDRPNCLNADPCDPATAPKNKSTKPLDALV 308
QY 531 -----TFKDG-----ALALSCMADYGAHPVPFPFILAIGYKKG 562
Db 309 LNGGCGQLSRKYMHWQBEELIVGTIVKNATGKLVSAHALQTMFLQMTPKQMYHFERY--365
QY 563 DYSEAEALIMTFLSNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMFQVTFTAERSLE 622
Db 366 DY-----VSHINWE-----DRAAILEAWQRTYVEVHQSVAPNSTQKVLFPFTT-TTLD 414
QY 623 DEINRTTAEPLIPATSYIVIFYISLALGSYSWSRVNVDISKATIGLGGVAVVLGAVMA 682
Db 415 DILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKSGAVGLAGVLLVALSVAA 469
QY 683 AMGFYSYLGIRSSVLQVVPFLVSVGADNIFIVLEYORLPRRPGEPREVIHGAQR 742
Db 470 GLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLLAHAPSETGQNKRIIPEDRTGBCLXR 529
QY 743 VAPSMILCSLSEAICFFLGNLTMPNAVRTFALTSLAVILDPQLQMSAPVALLSLDSKQ 802
Db 530 TGASVALTISINVTAPFMAALIPALRAFSLQAQAVVVFNFAMVLLIPAILSLMDLYR 589
QY 803 EASRLDVCC-----VKQEL-----PPQCGEGL-----827
Db 590 EDRLDIFCCPTSPCVSRVIOVEPOAYTEPHNTRYSPPPPYTSHSPAHEITHITMQSTVQ 649
QY 828 -----827
Db 650 LRTEYDPHTVYVYTTAEPRSEISVQPVTVTQDNLSQSPSESTSTRDLLSQFSDSLHCL 709
QY 828 -----LLGPFQKAYAPFLHWTGRGVLLFLALQVSLYSMSCHISVGLDQELALPK 879
Db 710 BPCCTWTLSSFAEKHYAPFLLLKPAKVVVILLFGLLGVSLYGTTRVRDGLDLTDIVR 769
QY 880 DSYLDYFLNRYEFGVAPVYVFTTLGYNFSSRAGMAICSSAGCNSFTQKIQYAT- 938
Db 770 ETRYDFIAAQKYPF-----YNM-----YIVTKADYPNI 801
QY 939 -----EPFESYLAIPASS-----WVDDFIDWL-----TPSS-----965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQWMLHYFRDMLQGLQDAPDSDMETGRIMPNNYKN 861
QY 966 -----CCRLYISGPNKDKPCPSTVNSLNCNCKMSITMGSVRSPVSEQPHKYL-PWFLN 1017
Db 862 GSDGVLAYKLLVQTGRDK--PIDISQLT--KQRLVDADGIIINPSA--FYIYLTAWSN 915
QY 1018 D-----RPNIKCPKGLAAY--STSVNLTSQGVLASRPMAYHKPLKNSQD 1061
Db 916 DPVAYAAQANIRPHRPWVHDK---ADYMETRLRIPAAEPIEYAOQPPFYINGLRDTS 972
QY 1062 YTEALRAARELAANITA-DLRKVPCTDPAFEVPPYITNVFYEQYLTILPEGLFMLSCL 1120
Db 973 FVEATEKRVICNNYTSLSGSSYPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIYVILVDTVGFMAWLDISYNAVSLINLSAVGMSVE 1180
Db 1023 ACTFLVCAVFLNPTAGII-VNVLALMTVELFGWMLGIGIKLSAVPVVILLASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVFAVAVMTNLPGLILVLGLAKAQIQLF 1238
Db 1082 FTVHVAFALTAIGDKN---HRAMLALEHMFAPVLDG-AVSTLLGLVLMLAGSEFDFIVRY 1137
QY 1239 FFRNLNLTLLGLLHGLVFLPVLISYVG--PDVNPALALEQ---KRAEBAVAAMVASC 1293

Db 1138 PFAVLAITLVGLNGLVLLPVLLSFFGCPCEVSPANGNLRLPTSPPEPPSVVREAVPP 1197
Qy 1294 NHPSRVS-TADNIYNHSPFESI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISBELRQYEAQOGAG 1234

RESULT 19
US-08-954-668-10
; Sequence 10, Application US/08954668
; Patent No. 6172200
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,668
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-668-10

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

Qy 356 LVLSVIPVVALAAGLVFTEITDDELWSAPNSQARSEKAFHQHGFPPFRN-QVILPA 414
Db 88 LVVGLLIFGAFVGLKAALETNVEELMVEGVGRSRELYTRQKIGESAMFNPLMIQT 147
Qy 415 PNRSSRYDLSLLGPNFSGIOLDLLELLE--LQERLRLHLOVQWSPFQARNISLODICY 472
Db 148 PKEEG-----ANVLTEALLQHLDSALQASRVHVVMYN-----RQWKLHLCY 190
Qy 473 AP-----LNPDTSLVDCCINSLLQVFQNNRTLLLTANQTLMO----- 512
Db 191 KSGELITETGYMDQITBYLPCLITPLDCEWGAQLQSGTA--YLLGKPLRWTFNFDPL 248
Qy 513 -----TSQVDWKDHFLY-----CAN-----APL----- 530
Db 249 EFLEELKKINYQVDSWEMLNKAEGVGYMDRFLCNPADPCPATPNKSTKPLDVALV 308
Qy 531 -----TFKDG-----ALALSCMADYGAPVFPFLAIGYKKG 562
Db 309 LNGGCGLSRKYMHQBELIVGGTVKNATGKLVSAHALQTMFQMLMTPKQMYEHRGY--- 365

Qy 563 DYSEAEALIMTFGLSNYPAGDPRPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTFPAERSLE 622
Db 366 DY-----VSHINWNE-----DRAAAILLEAQRTYVEVVHQSVAFNSQKVLPTT-ITLD 414
Qy 623 DEINRTAEDLPFATSYIVIFLYISALGSYSWSRWVDSKATLGLGGVAVVLGAVMA 682
Db 415 DILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKSGAVGLAGVLLVALSVAA 469
Qy 683 AMGFPSYLGIRSSILVLOVVPFELVSVGADNIFVLEVQRLPRRCPPEPEVHIGRALGR 742
Db 470 GLGLCSLIGISFNAATQTVLFFLAGVGVDVDFLLAHAFSETGQNKPIPEDRTRGECLKR 529
Qy 743 VAPSMLLCSLSEACFFELGALTTPMPAVRTFALTSLGLAVILDFLQMSAFVALLSLSKQ 802
Db 530 TGAVALTSISNVTAFPMALIPALRAFSLQAAVVVFNFAMVLLIFAILSMDLVRR 589
Qy 803 EASLDVCCC-----VRQDEL-----PPGQGEGL----- 827
Db 590 EDRLDIFCCFTPCVSRVQVEPQAYTEPHSNTRYSPPPPTSHSFAHETHITMQSTVQ 649
Qy 828 ----- 827
Db 650 LRTEYDPHTHVYTTAEPRSEISVQPVTVTDQNLSCQSPSTSTRDLLSQFSDSLHCL 709
Qy 828 -----LLGFFQKAYAPFLLHWITRGVLLFLALFGVLSYXSMCHISVGLDQELALPK 879
Db 710 EPPCTKWTLSFAEKHYAPFLLKPAKVVVILLFLGLLSVLYGTTRVRDGLDLDITVPR 769
Qy 880 DSYLLDYFLNRYFEVGAPVYVTTILGYNFSSEAGMNAICSSAGCNCNFSFTQKIQVAT- 938
Db 770 ETREYDFIAAQKYFSF-----YNN-----YIVTQADYDNI 801
Qy 939 -----EPPEQSILAIPASS-----WVDDFIDWL-----TPSS--- 965
Db 802 QHLLYDLHKFSNVKYVMLEENKQLPQWLHYFRDMLQLODAFDSQWETGRIMPNYKN 861
Qy 966 -----CCRLYISGPNKDFCPSVTNSLNCMLKMSITMGSVRPSVQFHKYL-PWFLN 1017
Db 862 GSDDGVLAYKLLVOTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAMVSN 915
Qy 1018 D-----RPNIKCPKGGLAAY--STSVNLTSDQVLAERFMAHKLKNSQD 1061
Db 916 DPVAYAASQANIRPHRPEWHDK---ADYMPETRLRIPAAEPIEYQAQPFYINGLURDTS 972
Qy 1062 YTEALRAARELAANITA-DLRKVPCTDPAFVFPFYITITNVFYEQYVLTILPEGLFMLSCL 1120
Db 973 FVEAIEKRVVICNNVTSGLSSYPNG-----YFP-----LFEQYISLRHWLLLSISVVL 1022
Qy 1121 VPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGPMLWDISYNVSLINLVSAYGMSVE 1180
Db 1023 ACTFLVCVAVFLNPNWTAGII-VMVILAMTVLFGMMLGIGIKLSAVPVVILIASVIGVE 1081
Qy 1181 FVSHITRSP--AISTKPTWLERAKETISMGSAVPAGVAMTNLPGLVILGLAKAQLIQIF 1238
Db 1082 FTVHVALAFLTAIGDKN---HRAMLALEHMFAPVLDG-AVSTLLGLVLMAGEFPFIVRY 1137
Qy 1239 PPRMLLTLLGLLHGLVFLPVLSYVG--PDVNPALALEQ---KRAEEAANAAMVASCP 1293
Db 1138 FPAVALITLVGLNGLVLLPVLLSFFGCPCEVSPANGNLRLPTSPPEPPSVVREAVPP 1197
Qy 1294 NHPSRVS-TADNIYNHSPFESI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISBELRQYEAQOGAG 1234

RESULT 20
US-08-918-658-10
; Sequence 10, Application US/08918658
; Patent No. 6429354
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L

TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICANT: US/08/918,658

FILING DATE: 22-Aug-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/656,055

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/540,406

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Bertram I

REGISTRATION NUMBER: 20015

REFERENCE/DOCKET NUMBER: a60190-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1434 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-08-918-658-10

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 3.3e-51;
Matches 283; Conservativity 171; Mismatches 420; Indels 363; Gaps 44;

QY 356 LVLSVIPVVALAAGLVFTLTTPDELWSAPNSQARSEKAFHDQHPFPFRTN-QVILTA 414
DB 88 LVGLLIFGAPAVGLKAANLETVNELVVEGVRSRELYTRQKIGBEAMFNPQMLOT 147
QY 415 PNRSSRYDSLLGLGPKNFGIIDLLELLE--LQERLRHLQVMSPEAQRNLSQDICY 472
DB 148 PKBEG-----ANVLTTEALLQLHDSALQASRVHVMYN---RQWGLEHLCY 190
QY 473 AP-----LNPDNTSLYDCINSLLQYFQNNRTLLILLTANQTLMOQ----- 512
DB 191 KSGELITETGYMDQITFEIYLYPLCIITPLDCPMEGAKLQSGTA--YLLGKPLRLWTFNFDPL 248
QY 513 -----TSQVDKDHFLY-----CAN-----APL----- 530
DB 249 EFLBELKKINYQVDSWEEMLNKAEVGHGYNDRPCLNPADPCPATAPNKNSTKPLDVALV 308
QY 531 -----TFKDGTT-----ALALSCMADYGAPVFPFLAIGYKKG 562
DB 309 LNGGCGQLSRKYMHWQDEELIVGTVKNATGKLVSAHALQTMFQMTFKQMYEHFRGY--- 365
QY 563 DYSEBALMTSLNYPAGDPRLAQAKLWEBAFLMEARAFORMMAGMFOVTFATERSLE 622
DB 366 DY-----VSHINWNE-----DRAAAITLEAQRITYEVVHQSVAPNSTQKVLPEFTT-TTLD 414
QY 623 DEINRTTADLPFATSYIVIFLYISLALGSYSSNSRVVMDSKATIGLGGVAVLGAWMA 682
DB 415 DILKPSDVSVIRVASGYLLMLAYACLTW---LRMD--CSKSGQAGLAGVALLVALSVAA 469
QY 683 AMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRALGR 742

DB 470 GLGLCSLIGISFNAATTQVLPLFALGVGDVDFELLAFAFSETQCNKRIPEDRTGCLCKR 529
QY 743 VAPSMILCSLSSEALCFPLGALTMPAVRTFALTSLGLVDFLLQMSAFVALLSLDSKQ 802
DB 530 TGAVALTSISNVTAFMAALPIPALRAFSQAQAVVVVFNFAWVLLIPAILSLMDLYR 589
QY 803 EASRLDVCCC-----VKQBEL-----PPPGQGEGL----- 827
DB 590 EDRDLIFCCFTSPCVSRVIQVEPQAYTEPHSNTRYSPPPPYTSHSFAHETHITMQSTVQ 649
QY 828 ----- 827
DB 650 LRTEYDPHTHYVYTTAEPRSEISVOPVTVDNLSCQSPSTSTRLLSQFSSDSLHCL 709
QY 828 -----LLGFFOKAYAPFLHMTITGVVLLFLALFGVLSYSMCHISVGLDQELALPK 879
DB 710 EPPCTWTLSSFAEKHYAPFLKPKAKVVLILLGLLGVSLYGTTRVRDGLDLDIVPR 769
QY 880 DSYLLDYFLNRYFEVGPVYFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYAT- 938
DB 770 ETREYDFIAAQPKYFSF-----YNM-----YIVTQKADYPMI 801
QY 939 -----EPPEQSYLAIPASS-----WDDPIDWL-----TPSS--- 965
DB 802 QHLLYDLHKFSNVKYVMLEENKQLPQMWLHYFRDWLQGLQDAFSDSDWETGRIMPNNYKN 861
QY 966 -----CCRLVISGPNKDKCPCSTVNSLNCILKNCMSITWGSVRPSVEQPHKYL-PWFLN 1017
DB 862 GSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVDADGIINPSA--FYIYLTAWVN 915
QY 1018 D-----RPNIKCPKGLAAY--STSVNLTSQGVGLASRFMAYHKPLKNSQD 1061
DB 916 DPVAYAASQANIRPHRPEWVHDK---ADYMPETRLRIPAAEPIEYAFQFFYLLNGLEDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPGTDPAPFVFPYTTINVFYEQYLTILPGLFMLSLCL 1120
DB 973 FVEATEKVRVICNNVTSLGLSSYPNG-----YFP-----LPWEQYISLRHMLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLLSIVMLIVDTGVFMALWDISYNAVSLNLSAVGMSVE 1180
DB 1023 ACTFLVCVAVFLNPNWTAGII-VMLALMTVELFGMGLGKLSAVPVVVLVTSVIGIVE 1081
QY 1181 FVSHITRSF--AISTKPTWLERAKEATISMGSAGVAGVAMTNLPGLILVLGAKAQLQIF 1238
DB 1082 FTVHVALAFLTAIGDKN---HRMLALEHMPAPVLDG-AVSTLLGVMLAGSEFDFIVRY 1137
QY 1239 PFRMLLITLLGLHLVPLVILSYVG--PDVNPALALEQ---KRAEBAVAAMVVASCP 1293
DB 1138 PFAVLAITLVGLVGLVLLPVLSSFFGCPCEVSPANGLNRLTPSPBPPPSVVRFAVPP 1197
QY 1294 NHPSRVS-TADNIYVNHSPGSI-----KGAG 1319
DB 1198 GHTNNGSDSDSEYSSQTTVSGISEELRQYEAQOAG 1234

RESULT 21

US-09-724-631-10
Sequence 10, Application US/09724631
Patent No. 6551782
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
JOHNSON, RONALD L
GOODRICH, LISA V
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:	
MEDIUM TYPE:	Floppy disk
COMPUTER:	IBM PC compatible
OPERATING SYSTEM:	PC-DOS/MS-DOS
SOFTWARE:	Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:	
APPLICATION NUMBER:	US/09/724,631
FILING DATE:	28-NO. 6551782-2000
CLASSIFICATION:	<Unknown>
PRIOR APPLICATION DATA:	
APPLICATION NUMBER:	US/08/656,055
FILING DATE:	1996-05-31
APPLICATION NUMBER:	08/540,406
FILING DATE:	<Unknown>
ATTORNEY/AGENT INFORMATION:	
NAME:	Rowland, Bertram I
REGISTRATION NUMBER:	20015
REFERENCE/DOCKET NUMBER:	a60190-1
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	415-781-1989
TELEFAX:	415-398-3249
INFORMATION FOR SEQ ID NO: 10:	
SEQUENCE CHARACTERISTICS:	
LENGTH:	1434 amino acids
TYPE:	amino acid
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE: protein	
SEQUENCE DESCRIPTION: SEQ ID NO: 10:	
US-09-724-631-10	
Query Match	
Best Local Similarity 9.4%; Score 646.5; DB 2; Length 1434;	
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;	
QY	356 LVLSVTPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFDQHFQPPFRN-QVILTA 414
DB	88 LVGLIFGAFAGVLAAGLVFTLTTPDVELMSAPNSQARSEKAFDQHFQPPFRN-QVILTA 414
QY	415 PNRSSRYSLIAGPKNFSGILDLILLLELLE--LQERLRHQLQVMSPEAQRNISLODICY 472
DB	148 PKEEG-----ANVLTEALLQHLDSALQASRVHVMYN-----RQWKEHLCY 190
QY	473 AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLILLTANQTLMGQ----- 512
DB	191 KSGELITETGYMDQIIIEYLIPCLIIITPLDCFMEGAKLQSGTA--YLLGKPLRLWNTNFDPL 248
QY	513 -----TSQVDWKDHFLLY-----CAN-----APL----- 530
DB	249 EFLEELKKINYQVDSWEMLNKAEGVGHGYMDRCLNPADPCPATAPKNKSTKPLDVALV 308
QY	531 -----TFKDG-----ALALSCMADYGAPVFPFFLAIGYKKG 562
DB	309 LGGCQGLSKRYMHQBEELVGTGVKNATGKLSAALQTMFQMTPKQMYEHFRGY--- 365
QY	563 DYSEAEALIMTFSINYPAGDPLAQAQLWEEAFLEEMAFQRRMAGMQVITFAERSLE 622
DB	366 DY-----VSHINWNE-----DRAAAILEAWQRTYVEVHQSVAPNSTQKVLPEPT--TTLD 414
QY	623 DEINRTTAEDLPFATSYVIFLYISALGYSYSSWGRVMVDSKATGLGSGVAVVLGAVMA 682
DB	415 DILKPSDVSIVIRVASGYLLMLAYACLTM---LRWD---CSKSGAGVLAGVLLVALSVAA 469
QY	683 AMGFPSYLGIRSSLVILQVVPFVLVSVGADNIIFVLEYQRLPRRGPCEPREVHIGRALGR 742
DB	470 GLGLCSLIGISFNAATTVQLPFLALGVDDVDFLLAHAFSETGQNKRIIPEDRTGCLXR 529
QY	743 VASMLCSLSEALCFPLGALTWPVARTALTSLGLAVILDFLLQMSAFVALLSLDSKQ 802
DB	530 TGAASVALTISNTVTAFFMAALIPILPAURAFSLQAAVVVFNFAVLLIFFAILLSMDLYR 589
QY	803 EASRLDVCCC-----VKPQEL-----PPFGQGEGL----- 827

DB	590 EDRLDIFCCFTSPCVSRVQVPEQAYTBPHSNTRYSPPPPTSHSPAETHITMOSTVQ 649
QY	828 ----- 827
DB	650 LRTEYDPHTHYTTAEPREISVQPVTVTQDNLSCQSPSTSTRDLLSQFSSSLHCL 709
QY	828 -----LILGFFQKAYAPFLLHMITRGVVLLFLALFGVSLYSMCHSISVGLDQELAPK 879
DB	710 EPCTKWTLSFPAEKHYAPFLLKPAKVVILLFLGLGVSGLYGTTRVRDGLDLDIVPR 769
QY	880 DSYLLDYFLNRYFEVGAAPVYFVTTLGYNFSSEAGNAICSSAGACNFFFTQKIQAT- 938
DB	770 ETREYDFIAAQFKYFSP-----YNN-----YIVTKADYENI 801
QY	939 -----EPQSQSYLAIPASS-----WDDDFIDML-----TPSS--- 965
DB	802 QHLLYDLHKSFSNVKVMLEENKQLPQMWLHYFRDMLQGLQADAPDSWETGRINPNYKN 861
QY	966 -----CCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQHFHYL-PWFLN 1017
DB	862 GSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYILTAWVSN 915
QY	1018 D-----RNIKCPKGGLAAY--STSVNLTSQGVLASRFMAYHKPLKNSQD 1061
DB	916 DPVAYAAQANIRPHRPEWVHDK---ADYMPETRLRIPAAEPTEYAQFPFYINGLRDTS 972
QY	1062 YTEALRAARELANITA-DLRKVPGTDPAPFVPPYTTITNVFYOYLTILPEGLFMLSCL 1120
DB	973 FVEAIEKRVICNNYTSIGLSSYPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
QY	1121 VPTFAVSCLLGLDLRSGLLNLLSIVMLVDTVGFMAIDISYNVSLINLVSAVMSVE 1180
DB	1023 ACTLVCAVFLNPTAGII--VMVALMTVELFGMGLIGIKLSAVPVPVILIASVGIGVE 1081
QY	1181 FVSHITRSP--AISTKPTWLERAKETISMGSAVPAGVAMTNLPGILVLGLAKAQLQIF 1238
DB	1082 FTVHVALAFLTAIGDKN---HRAMLALEHMPAPVLDG--AVSTLLGLVLMLAGSEDFIVRY 1137
QY	1239 FFRNLNLTITLGLAGLVFLPVILSVYG--PDVNPALALEQ---KRAEEAFAVAVMVASCP 1293
DB	1138 FFAVALIUTVLGVNLGLVLLPVLISFPFGCPCEVSPANGNLRLPTSPSPFPVVRFAVPP 1197
QY	1294 NHPRSVS--TADNIYVNHSPFGSI-----KGAG 1319
DB	1198 GHTNNGSDSSDSEYSSQTTVSGISELRLQYEAQAG 1234
RESULT 22	
US-08-954-701A-10	
Sequence 10, Application US/08954701A	
Patent No. 6610507	
GENERAL INFORMATION:	
APPLICANT: SCOTT, MATHEW P	
APPLICANT: GOODRICH, LISA V	
APPLICANT: JOHNSON, RONALD L	
TITLE OF INVENTION: Patched Genes and their Use	
NUMBER OF SEQUENCES: 19	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Foley, Hoag & Eliot	
STREET: One Post Office Square	
CITY: Boston	
STATE: MA	
COUNTRY: USA	
ZIP: 02109	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Floppy disk	
COMPUTER: IBM PC compatible	
OPERATING SYSTEM: PC-DOS/MS-DOS	
SOFTWARE: ASCII(text)	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/08/954,701A	
FILING DATE: 20-OCT-1997	
CLASSIFICATION: 435	

		TOPOLOGY: linear		
		MOLECULE TYPE: protein		
		SEQUENCE DESCRIPTION: SEQ ID NO: 10:		
		US-09-754-032-10		
		Query Match		9.4%; Score 646.5; DB 2; Length 1434;
		Best Local Similarity		22.9%; Pred. No. 3.3e-51;
		Matches		283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
Qy	356	LVLVSVVVAALAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFPFRTN-QVILTA	414	
Db	88	LVVGLLIIFGAFVGLKAANLETNVEELWVEGVRVSRRELNTRQKIGEEAMFNPLMIQT	147	
Qy	415	PNRSRYRYSLLGPKNFSGILDLLLELLE--LQERLHLQVWSPQAQRNLSLDICY	472	
Db	148	PKEEG-----ANVLTTEALLQHLDSALQASRVHYMYN-----RQWLEHLCY	190	
Qy	473	AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQ-----	512	
Db	191	KSGELITETGYMDQIIEVLYPCLIIITPLDCFWEGAKLOSGETA--YLLGKPLRWTFDPL	248	
Qy	513	-----TSQVMDKDHFLY-----CAN-----APL-----	530	
Db	249	EFLEELKKINYQVDSWEEMLNKAEGVGHGYMDRPLNADPCPATAPNKNSTKPLDVALV	308	
Qy	531	-----TFKDGTT-----ALALSCMADYAGVPFPFLAIGYKKG	562	
Db	309	LNGGCCGLSRKYMHWQEBELIVGGTVKNATGKLVSAHALQTMFQMTPKQMYEHFRGY---	365	
Qy	563	DYSEAEALIMTFSLNYPAGDPRPRLAQAKLWEEAFLEEMRAFORRMAGMFQVTTAERSLE	622	
Db	366	DY-----VSHINWNE-----DRAAILEAWQRTYVEVHQSVAPNSTOKVLPFTT-TTLD	414	
Qy	623	DEINRTABDLPIFATSYIVIFLYISLALGSYSWSRVNVDKATLGLGSAVVLGAVMA	682	
Db	415	DILKSFSDSVIRVASGYLLMLAYCLTW---LRWD--CSKQGAAGVLAGVLLVALSVAA	469	
Qy	683	AMGFFSYLGRSLVTLQVVPFLVLSVGADNIFIVLEYQRLPRRPEPREVHIGRALGR	742	
Db	470	GLGLCSLIGISFNAATQVLPFLALGVGDVDFLLAHAFSETGQNKRIPEFDTGECLKR	529	
Qy	743	VAPSMLLCSLSBAICFFGALTPMPAVRTFALTGSLAVILDFLQMSAFVALLSDSKRQ	802	
Db	530	TGASVALTSISNVTAFMAALIPALRAFSLOAAVVVFNFAWVLLIFPAILSMWLYR	589	
Qy	803	EASRLDVCCC-----VKPQEL-----PPGQGEGL-----	827	
Db	590	EDRLDIFCCFTSPCVSRVQVEPQAYTEPHSNTRYSPPPVYTSFSAHETHITMQSTVQ	649	
Qy	828	-----	827	
Db	650	LRTEYDPHTHYVTTAEPRESEISVQPTVTQDNLSCQSPSTSTRDILLSQFSDSLHCL	709	
Qy	828	-----LLGFQKAYAPFLHWITRGVLLLFALFGVSLYSMSCHISVGLDQELALPK	879	
Db	710	EPCTKWTLSSFAEKYAPFLPKAKVVILLFLGLGLVSLYGTTRVDRGLDLDIVER	769	
Qy	880	DSYLLDYFLNRYREVGAQVYFVTTLGYNFSSAGHNAICSSAGCNFSTFKIQYAT-	938	
Db	770	ETREYDFIAQPKYFSF-----YNM-----YIVTQADYPNI	801	
Qy	939	-----BFPQSYLAIPASS-----WVDDFDIML-----TPSS---	965	
Db	802	QHLLYDLHKFSFNKYNKYLEENKQLPQMWLHYFRDWLQGLQADAFDSWETGRIMPNNYN	861	
Qy	966	-----CCRLYISGPNKDKFCPSTVNSLNCNCKMSITMGVSRPSVEQEHKL-PWFLN	1017	
Db	862	GSDGVLAYKLLVQTGRDK--PIDISQIT--KQRLVDADGIINPSA--FYIYLTAWSN	915	
Qy	1018	D-----RPNIKCPKGLAAY--STSUNLTSQGVQVLAIRFMAHKLKNSQD	1061	
Db	916	DPVAYAAQANIRPHREPWVHDK--ADYMPETRLRIPAAEPIEYAFQFPFYLNGLRDTS	972	

Db 366 DY-----VSHINWE-----DRAAILEAWQRTYEVVHOSVAPNPTQKVLPETT-TTLD 414
QY 623 DEINRTTAEDLPFATSYIYIFLYISALGSSYSSWRVMVDKATLGLGGVAVVLGAVMA 682
Db 415 DILKSFSDSVIRVASSYLLMLAYACTM--LRWD--CSKSQAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVILQVPPFLVSLVGADNIPIFVLEYQRLPRRGPBPREVHIGALGR 742
Db 470 GLGLCSLIGISFNAATTVLPFLALGVGVDDVFLAHAFSETQNKRIPFEDRTGECLAR 529
QY 743 VAPSMMLCSLSEALCFGLGALTMPAVRTFALTSLGLAVILDLLQMSAFVALLSDSKRO 802
Db 530 TGSVALTSISNTAFMAALIPALRAFSLQAAVVVFNFAVLLIPAILSDMLYXR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPGQGEGL-----827
Db 590 EDRRLDIFCCFTSPCVSRVIOVEPQAYTEPHSNTRYSPPPYTHSHFAHETHITMQSTVQ 649
QY 828 -----827
Db 650 LRTEYDPHTVHYVYTTAEPRSEISQVPTVTQDNLSQSPSESTSTRDLSQFSDSLHCL 709
QY 828 -----LLGFPQKAYAPFLHMITRGVVLILFLALFGVSLYSXSMCHISVGLDOELALPK 879
Db 710 EPPCTKWTLSSEAHHYAPFLPKAKVGVVILLFLGLLGVSLYGTTRVRDGLDLDIVPR 769
QY 880 DSYLLDYFLNRYFVPGVAVPVVTVTLGYNFSEAGMNAICSSAGCNFSFTQKIQYAT- 938
Db 770 ETRYDFIAAQPKYFSF-----YNM-----YIVTQKADYPMI 801
QY 939 -----EFPEQSYLAI PASS-----WDDPIDWL-----TPSS-----965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQMWLHYFRDMLQGLQDAFSDWETGRIMPNNYKN 861
QY 966 -----CCRLYISGNKDKFCPSYVNSLNCNCHSIYMGSVRPSVEQPHKYL-PWFLN 1017
Db 862 GSDGVLYAKLLVQTSRDK--PIDISQLT--KQRLVDADGGINPSA--FYLYLTAWVSN 915
QY 1018 D-----RPNIKCPKGGAAAY--STSVNLTSDBGVLAASRFMAYHKPLKNSOD 1061
Db 916 DPVAVASQANTPRPPEWHDK--ADYMPETRIIPAEPIEYIAQFFPYLNGRLDSD 972
QY 1062 YTEALRAARELANITA-DLRKVPGTDPAPFEPVPTITNVFYEQLTILPEGLFMLSCL 1120
Db 973 FVEATEKRVICNNVYTSGLSVPNG-----YFP-----LFWEQYISLRHMLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLBSGLNLLSIVMILVDTVGPMALWDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCAVFLNPTWAGII-VNVALMTVELFGMMGLIGIKLSAVPVVILLIASVGIGVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIF 1238
Db 1082 FTVHVVALAFLTAIGDN-----HRAMLAEHMFAPVLDG-AVSTLLGVLMLAGSEFPFIVRY 1137
QY 1239 PFRANLLITLLGLHGLFPLVLSYVG--PDVNPALAEQ---KRAEBAVAANVASC 1293
Db 1138 PFAVLAAILTVLGLNGLVLLPVLLSFFGCPCEVSPANGNLRLTPSPPEPPSVVRFVPP 1197
QY 1294 NHPRSVS-TADNIYVNHSPGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISBELROYEAQOAG 1234

RESULT 25

PCT-US95-13233-10

; Sequence 10, Application PC/TUS9513233
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13233
; FILING DATE: 06-OCT-1990
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-13233-10

Query Match 9.4%; Score 646.5; DB 4; Length 1434;

Best Local Similarity 22.9%; Pred. No. 3.3e-51;

Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

QY 356 LVLSVIPVVALAAGLVFTLTDPVLSWAPNSQARSEKAFHQHGFPPFRN-QVILFA 414
Db 88 LVVGLLIIFGAFVGLKAANLETNVEBELVVEVGRVSRRLNYYTRQKIGEBAMFNQMIQT 147
QY 415 PNRSSRYVDSLGLGPKNFSGILDLLLELLS--LQERLRHLQVNSPEAQNISLQDICY 472
Db 148 PKBEG-----ANVTTEALQHLDSALQASRVHVMTYN-----RQWKLEHLCY 190
QY 473 AP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTANQITLMGO-----512
Db 191 KSGELITETGYMDQIIIEVLYPCLLIITPLDCFWEGAKLQSGTA--YLLGKPLRWTFDPL 248
QY 513 -----TSQVDWKDHFY-----CAN-----APL-----530
Db 249 EFLEBLKKINYQVDSWBEMLNKAARYGHGYMDRPLNPDPCPATAPNKNSTKPLDVALV 308
QY 531 -----TFKQGT-----ALALSCMADYGA PVPPFLAIGYKQK 562
Db 309 LNGGCGQLSRKYMHWQEBELIVGGTVKNATGKLVSAHALQTMPLQMPKQMYEHFRGY---365
QY 563 DYSEALALIMTFSLNNYPAGDPRLAQAQKLWBEAFLEEMRAFORRMAGMFQVTFTRAESLIE 622
Db 366 DY-----VSHINWE-----DRAAILEAWQRTYEVVHOSVAPNPTQKVLPETT-TTLD 414
QY 623 DEINRTTAEDLPFATSYIYIFLYISALGSSYSSWRVMVDKATLGLGGVAVVLGAVMA 682
Db 415 DILKSFSDSVIRVASSYLLMLAYACTM--LRWD--CSKSQAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVILQVPPFLVSLVGADNIPIFVLEYQRLPRRGPBPREVHIGALGR 742
Db 470 GLGLCSLIGISFNAATTVLPFLALGVGVDDVFLAHAFSETQNKRIPFEDRTGECLAR 529
QY 743 VAPSMMLCSLSEALCFGLGALTMPAVRTFALTSLGLAVILDLLQMSAFVALLSDSKRO 802
Db 530 TGSVALTSISNTAFMAALIPALRAFSLQAAVVVFNFAVLLIPAILSDMLYXR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPGQGEGL-----827
Db 590 EDRRLDIFCCFTSPCVSRVIOVEPQAYTEPHSNTRYSPPPYTHSHFAHETHITMQSTVQ 649

```
QY 828 ----- 827
Db 650 LRTEYDPTHVYTTABRSEISVQPVTVTDNLSCQSPSTSTRDQLSQFSSLHCL 709
QY 828 -----LLGFFOKAVAPFLHWTIRGVLLFLALFGVSLYGMCHISVGLDQELALPK 879
Db 710 EPPCTKWTLSFAEKHAPFLKPKAKVWVILLFGLGLVSLYTTVRDGLDLDIVPR 769
QY 880 DSYLLDYFLNRYFEVGAPVYFVTTGLYNNFSSBAGMNAICSSAGCNFFSTQKIQYAT- 938
Db 770 ETVEDFTAAQKIFYSP-----YXN-----YIVTKADYPNI 801
QY 939 -----EPQESYLAIPASS-----WUDDFDLW-----TPSS--- 965
Db 802 QHLLYDLHKFSNRYKVMLEENKQLPQMWLFYFDWLQGLQADPDSWETGRIMPNNYKN 861
QY 966 -----CCRLYISGPNKDFCPSVNSLNCNKCMSITMGSVRPSVEQPHKYL-PWPLN 1017
Db 862 GSDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWSN 915
QY 1018 D-----RPNIKCPKGLAAY--STSVNLTSQGVLASRPMAYHKPLKNSQD 1061
Db 916 DPVAYAAQANIRPHRPEWVHDK--ADYMPETRLRIPAAEPYEAQPPFYLNGLRDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPCTDPAFVFPYTTINVEYQYLTILPGLFMLSCL 1120
Db 973 FVEATEKVRVCNNYTSLSGLSSYPNG-----YFP-----LFWEQYISLRHLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLSGLLNLSTVMILVDTPGPMALWDISYNNAVSLINIVSAGMSVE 1180
Db 1023 ACTFLVCVFLNPNWTAGII-VWULAMTWELFGMGLIGIKLSNVPVILLIASVGIGVE 1081
QY 1181 FVSHITRSF--AISTKPTWLERAKETISMGSAVEPAGVAMTNLPGLVLGLAKAQLIOIF 1238
Db 1082 FTVHVALAFLTAIGDKN--HRAMLAEHMAPVLGD-AVSTLLGLVLMAGSEDFIVRY 1137
QY 1239 PPRMLLTLLGLLHGLVFLPILSYVG--PDVNPALAEQ---KRAEVAANVAVASCP 1293
Db 1138 FPAVLAITLVGLNGLVLLPVLSSFFGFCPEVSPANGLNRLFTSPSPFPVVRFAVPP 1197
QY 1294 NHPRSVS-TADNIYVNHSPFSGI-----KGAG 1319
Db 1198 GHTNGSDSDSEYSSQTTVSIGISELRQYEAQQAG 1234

RESULT 26
US-09-248-796A-16161
; Sequence 16161, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16161
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16161

Query Match 9.3%; Score 639.5; DB 2; Length 642;
Best Local Similarity 26.4%; Pred. No. 3.7e-51;
Matches 188; Conservative 123; Mismatches 291; Indels 111; Gaps 21;
QY 6 LRGWLLWALLLRLLAQSEPVTTIHQPQYCAFYDCGKNPGLSGSLMTLSNVSCLSNTPARK 65
```

```
Db 12 MRAICLLLLTINLAIAS-VSLSHKPGYCNTRYGNCGKKSVEGKPL-----PCAEEFPAVK 64
QY 66 ITGDHILLLOKICPRIPTYGNTQACCSAKQVLSLEASLSITKALLTFCPCACSNFVNLC 125
Db 65 ASQESREKLKSCGKDF-----DYICCSPEQIDILELNLKRVDPDLTSSCPACRKNFYDFC 120
QY 126 HNTCSNQSOLFIVNTRVAQLGAGQAPAVAYEAFYQHSFAEQSYDSRSRVVPAAATLAV 185
Db 121 QFSCSPNESQFVEIITKETARDTGKEIVTEINQYVEPGMANQFPDSCKNVKFLATN--- 176
QY 186 GTMCGVYGSAALCNAQRLNFGQDTGN--GLAPLDITFHLLPEQAVGSGQPLNEGVARC 243
Db 177 GYAMDLLGGGAKNVQLKPLGDEKPLLLGGSPYQINEVYKLP--ETDSGLVLRNEPLRDC 234
QY 244 NESQGDVATCSCDCAASCPATAPQALDSTFVGOMPSGLVLIILCISVFAVVTTLLV 303
Db 235 NDKE-----YKCACTDCBESCPKLPKADLTCKCTGVGLPCFSFSIII---IWSMIVLLG 287
QY 304 GERVAPAR-----DKSKVDPKKGTSLSDKLSFSTHTLLGQFFQGW----- 344
Db 288 GYHVLAKLKKERRRSTAESEDDSTMINPLFYAGLGKKRAKQFLSEIGLKIQDWFANI 347
QY 345 GTWVASWPLTILVLSVIPVVALAAGLVFTLTDPVELMSAPNSQARSEKAFHQHGFPP 404
Db 348 GYFCSKFPGISIGTSLAVVLLSLGLFLKQLETPVKLVSPNDPAYKNQOYFESFGEW 407
QY 405 FRTNOVLTAPNRSSRYDSLLGPKNFSGILDLDLLELLEQLERLRLHQVMSPEAQRN 464
Db 408 FRIEQVTVSS-----KDDGVLNWDIVKWMFDESKQLETL-----NEN 445
QY 465 ISLDQICVAPLNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFY 524
Db 446 VRLSDICFKPL--DET-----CALQSFQYFGD-----ISGLT-ETNWSKLQS 487
QY 525 CANAPLTFKGTALALSCMADYGPVFPFFLAIGYKDKYSEALIMTFSLNNYPAGDP 584
Db 488 CVDSF-----VNCLPTFQQLKPNIL---FDSNDISQAKAFTTVLVNSDTQNEH 534
QY 585 RLAQAKLWEERAFLEEMAFQRRMAGMQVTTAERSLEDEINRTTAEPLPIFATSYIVIF 644
Db 535 YTSNTISYEHLFKQWAADLQTEYNL-NIAYSTISLKEELNQSNTDIKTIAISLYWVF 593
QY 645 LYISLALGS-----YSSMSRVNVDSKATLGLGGVAVVLGAVMAAMGFFSYL 690
Db 594 IYASLALGGKLPNSANLYS-----LVKTRFTLGFSSIIILLSVTASAGSFLL 641

RESULT 27
US-08-857-636-60
; Sequence 60, Application US/08857636
; Patent No. 6552181
; GENERAL INFORMATION:
; APPLICANT: Dean, Michael Carlton
; APPLICANT: Hahn, Heidi Eve
; APPLICANT: Wicking, Carol
; APPLICANT: Christiansen, Jeffrey G.
; APPLICANT: Zaphiropoulos, Peter G.
; APPLICANT: Gailani, Mae R.
; APPLICANT: Shanley, Susan Mary
; APPLICANT: Chidambaram, Abirami
; APPLICANT: Vorechovsky, Igor
; APPLICANT: Holmberg-Lindstrom, Erika
; APPLICANT: Unden, Anne Birgitte
; APPLICANT: Gillies, Susan Alana
; APPLICANT: Negus, Kylie
; APPLICANT: Smyth, Ian Mcleod
; APPLICANT: Pressman, Carol Leah
; APPLICANT: Leffell, David J.
; APPLICANT: Gerrard, Bernard
; APPLICANT: Goldstein, Alisa Miriam
; APPLICANT: Wainwright, Brandon
; APPLICANT: Toftgaard, Rune Carl-Magnus
```

APPLICANT: Chenevix-Trench, Georgia
APPLICANT: Bale, Allen E.
TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSES: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/857,636
FILING DATE: 16-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,906
FILING DATE: 17-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P00011
FILING DATE: 21-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P00363
FILING DATE: 07-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/019,765
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-278200US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1296 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..1296
OTHER INFORMATION: /note= "amino acids encoded by human
OTHER INFORMATION: nevroid basal cell carcinoma syndrome
OTHER INFORMATION: (NBCCS) (PATCHED (PTC)) CDNA"
US-08-857-636-60

Query Match 9.1%; Score 628; DB 2; Length 1296;
Best Local Similarity 23.6%; Pred. No. 1.6e-49;
Matches 254; Conservative 165; Mismatches 384; Indels 272; Gaps 39;

QY 426 LLG--PKNFGSLDLDLLELLELLELLELHLOVMSPEAQRNLS-----LQDICVAPLNP 477
DB 97 LLGKPLRLWTPDLEFLEELKKI-----NQVDSWEEMLNKAEBVGHGMDRCPCLNPADP 151
QY 478 DNTSLYDCINSLLOYFQNNRTLLLTANTQIMLGQTSQ-VQWKOHFLYCANAPLTFKDG 536
DB 152 D-----CPATAPKNSTKPLDMALVNLGGCHLSRKYMHWQBELIVGG-----TVKST 200
QY 537 -----ALALSCWADYGAVPFPPLTGGYKGDYSEARALIMTFSLNNTYPAGDPRLAQAKL 591
DB 201 GKLSAHALQTMFLQMTPKQMY-----EHPKGYEY-----VSHINWNE-----DKAAAIL 247
QY 592 WEEAFLSEERAFORRMAGHFQVTFPAERSLEDEINRTTAEDLPFATSYIVIFLYISAL 651
DB 248 WORTYVEVHOSVAQNSTQKVLSTFT-TTLDDILKSPSDSVIRVAGYLMLLAYACITM 306

QY 652 GSYSSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFPSYLGIRSSLVILQVVPFLVLSVGA 711
DB 307 ---LRWD--CSKSGAVAGLAVLLVALSVAAGLGLCSLIGISFNAATQVLPFLAGVGV 361
QY 712 DNIPIFVLEYQRLPRPGEPREPHVHIGRAIGRVAPVSMMLCSLSEATCFPGALATPMPAVRT 771
DB 362 DDVFLAHAFSETGQNKRIPEPDRTGECILKRTGASVALTSISNVTAFPMALIPALRA 421
QY 772 FALTSLGAVILDFLLQMSAFVALLSLDSKROEASRLDVCCC-----VKQEL-- 818
DB 422 FSLQAAVVVVFNFAMVLLIFPAILSMDLYRRDRREDIFCCFTSPCVSRVIOVEPOAYTD 481
QY 819 -----PPPGQGEGL----- 827
DB 482 THDNTRYSPPPPYSSHSPAHEQTITMQSTVQLRTBYDPTHVYVYTTAEPRSEISVQPVTV 541
QY 828 -----LIGFFOKAYAPFLHMITROV 848
DB 542 TQDTLSCQSPSTSTRDLSQFSDSLHCLPEPPCTKWTLSFAEKHVAFLFKPKRAKV 601
QY 849 VLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTILGY 908
DB 602 VIFFLGLLGSVLYGTTVRDGLDUTDIVPRETRYDFIAAQKPYFSP-----Y 650
QY 909 NFSSEAGMNAICSSAGCNFSPFTQKIYAT-----EPPEQSYLAIPASS-----W 953
DB 651 NM-----YIVTQADYPIQHLHYDLHRSFSNVKYVMBENKQLPKMW 693
QY 954 VDDFIDWL-----TPSS-----CCRLYISGPNKDKPCPSTVNSL 987
DB 694 LHYPRDWLQGLQDAFDSWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISOL 751
QY 988 NCLKNCSITMGSVRPSVEQPHKYL--PWFLND-----RNKICPKGLAAY 1032
DB 752 T--KORLVADAGIINPSA--FYILTAWVSDNPVAYASQANIRPHRPEWVHDK--ADY 804
QY 1033 --STSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKVPCTDPA 1089
DB 805 MPETRLRIPAAEPIEYAGFPFVNLGRLDTSDFVEAIEKVRTICSNYTSIGLSSYPNG--- 861
QY 1090 FEVPYTTITNVFYQYLTILPEGLPMLSLCVPTPAVSCLLILGLDLRGLMLLSIVML 1149
DB 862 ---YFP---LFWEQYIGLRHMLLIFISVLACTPLCAVFLNPNWTAGII--VMVALMT 913
QY 1150 VDTVGFMAIWDISYNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLERAKEATIS 1207
DB 914 VELFGMGLIGIKUSAVPEVVIILASVGIVGVEFTVHVALAFLTAISDKN---RRVLALEH 970
QY 1208 MGSVAFVAGVAMTNLPGLVLGLAKAQLTIQIFFRLNMLLITLIGLHGLVLPVILSYVG- 1266
DB 971 MPAPVLDG-AVSTLLGVLMAGSEPDFIVRYFFAVLAILTILGLVNLGLVLPVLSFPGP 1029
QY 1267 -PDVNPALAEQ---KRAEAAVAAVMVASCPNHPESRVSTADNIYVNHFSFGSIKG 1317
DB 1030 YPEVSPANGLNRLPTSPSPPPSVVRFPAMPCHTH--SGSDSDSEYSSQITVSG 1082

RESULT 28
US-09-293-505-2
; Sequence 2, Application US/09293505
; Patent No. 6348575
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 2

```
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-293-505-2

Query Match      8.5%; Score 584; DB 2; Length 1203;
Best Local Similarity 22.9%; Pred. No. 2.1e-45;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;

Qy 306 RVAPARDKSMVDPKKGTSLSDKL--SPSTHTLLGQFFQG----WGTVWASWPLTILVLS 359
Db 3 RSPPLRELPSPSTPTPARTAAPQILAGSLKAPLWLRAYFQGLLFSLGCGIQRHCGKVLFLG 62
Qy 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHFQ--PPFTNQVILTAPNR 417
Db 63 LLAFGALGLRMAIETNLEQLWVEGSRVSEQLHYTKELGEEAAYTSQMLIQTAQOE 122
Qy 418 SSRYVDSLLGPKNFSGILDLDLLELEQLERHLQVMSPEAQRNISLQDICY---AP 474
Db 123 GEN-----ILTPALG-----LHLQALTASKVQVSLYCKSWDLNKKCYKSGVP 166
Qy 475 LMPDN-----TSLYDCCINSLLQVFNQNRITLLLTANOTLMGQTSQVDWKD----- 520
Db 167 LIENGMIEMWIEKLPFCVILTDLDFWEGAK---LQGGSAVLPGRPDIQWTNLDPEQLLE 223
Qy 521 -----HFLYCANAPLTFKDGSTALALSCMADYGA 548
Db 224 ELGFPASLEGFRELDDKAQVQVYGRPCLPHPDDLHCPPSPANNHHSRQAPNVAAHLSGCG 283
Qy 549 PVPEF-----LAIGYKGDYSE---AEALIMTF-----SLNNYPAGDPRLAQAKL 591
Db 284 HGFSHKFMWQBELLLGGWARPQOELLRAEALQSTFLLMSPRQLYEHFRGDYQTHDIGW 343
Qy 592 WEEAFLEENRAPQRRMAGMFQVTTFAERSLE-DEINRTTAEPLPIFATS-----Y 640
Db 344 SEEQASTVLQWQRFPVQLAQELPENASQOIHFASSTITLDDI-LHAFSEVSAARVWGGY 402
Qy 641 IVIFLYISALGYSWSRWVDSKATGLGCVAVVLGNVMAAGFFSVGLIRSSILVILQ 700
Db 403 LLMLAYACVTM---LRWD---CAQSGSVGLAGVLLVALAVASGLGCLALLGITFNAATQ 457
Qy 701 VVPEFLVLSGADNIFLVEYQRLPRRPGEPREHIGRALGVAPSMILCSLSEACIFFL 760
Db 458 VLPFLALGIVDDVFL--LAHAFTALPGTPOERMEGCELQRTGTSVLTSLNNMAAFILM 515
Qy 761 GALTMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLSKQESRLDVCCC----- 812
Db 516 AALVPIPALRAFSLQAAIVVGCTFVAVMLVFPAILSLDRRRCQRLDVLCCFSSPSCSAQ 575
Qy 813 ---VKQEL-----PPPGQGEGL-- 827
Db 576 VIQILPQELGDDGTVPVGIHAHLTATVQAFTHCEASSQHVVTILPPQAHVLPFSDPLGSEL 635
Qy 828 -----LLG-----PFQKAYAPFLHMITRGVWLLPLA 855
Db 636 FSPGSTRDLGQEEETRQKACKSLPCARWNLHAHPARYQFAPLILQSHAKAIVLVFCA 695
Qy 856 LFGVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAIVYFVTTLGNFS-SEA 914
Db 696 LLGLSLYGATLVQDGLATDVVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDVAHSQR 754
Qy 915 GMAAICSSAGCNFNSTQKIQYATEFPEQSYLAIPAS-----SWDDDFIDWLTPSSCC--- 967
Db 755 AL-----FDLHQRFSSL-----KAVLPPPTAQPRTHLYRNWLOGIQIAFPQ 798
Qy 968 -----RLYISGPNKOKFCPSTVNSLNCNKMCSITWGSVRPSVEQFHKYLPWFLND 1018
Db 799 DWASGRITRHSYNG-SED-----GALAYKLLIQTDGAQEPLD-FSQTITRKLVD 846
Qy 1019 RPNKCPKGLAAYSTSNITSDGVLAASFMAHYKP----- 1055
Db 847 REGLIIPPE--LFYMGTLVWVSSDPLGLAASQANFPYPPPEPWLHDKYDTTGENLRIPPAQ 904
```

```
Qy 1056 -----LKNQSDYTEALRAARELAANI-TADLRKVPQGTDPAPAFVFPYTTNVFY 1102
Db 905 LEPAQPPFLRLGLOKTDADFVEATEGARAAACAEAGAQGVHAYPSGSP------LFW 954
Qy 1103 EOYLITLPEGLFMLSIC--LVPTFVNSCLLLGLDLRSLNLLSIVMILYDVTGVMALWD 1160
Db 955 EOYLGL--RRCFLLAVCILLVCTFLVCALLLLNLPWTAGLI-VLVLAMMTVELFGIMFLG 1011
Qy 1161 ISYNAVSLINLVASGMSVEFVSHITRSPASTKPTWLERAKEATISMGSAVFAGVAMTN 1220
Db 1012 IKUSAIPVVTILVASVGIVFTVHVALGF-LTTQGSRNLRRAHALEHTFPVPTG-AIST 1069
Qy 1221 LPLGILVLGAKAQLIQIFFERLNLITLLGLLHGLVFLPVILSVGPDVNPALAEQRA 1280
Db 1070 LLGLLMLAGSHDFIVRYFFAALTVTLLGLLHGLVLLPVLSSILGPP--PEVIQMYKBS 1127
Qy 1281 EEAVAAV-----MVASCPNHPSPRSVSTADNIYN 1308
Db 1128 PELLSPPAQGGGLRWGASSSLPQSFARVTTSMVTVAIH 1165
```

RESULT 29

```
US-09-060-939A-2
; Sequence 2, Application US/09060939A
; Patent No. 6709838
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage, David A. Carpenter
; TITLE OF INVENTION: Patched-2
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,939A
; FILING DATE: 15-Apr-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-060-939A-2
```

```
Query Match      8.5%; Score 584; DB 2; Length 1203;
Best Local Similarity 22.9%; Pred. No. 2.1e-45;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
```

```
Qy 306 RVAPARDKSMVDPKKGTSLSDKL--SPSTHTLLGQFFQG----WGTVWASWPLTILVLS 359
Db 3 RSPPLRELPSPSTPTPARTAAPQILAGSLKAPLWLRAYFQGLLFSLGCGIQRHCGKVLFLG 62
Qy 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHFQ--PPFTNQVILTAPNR 417
Db 63 LLAFGALGLRMAIETNLEQLWVEGSRVSEQLHYTKELGEEAAYTSQMLIQTAQOE 122
Qy 418 SSRYVDSLLGPKNFSGILDLDLLELEQLERHLQVMSPEAQRNISLQDICY---AP 474
```

Db 123 GEN-----ILTPALG-----LHQAALTASKVQVSLYKGSWDLNKICYKSGVP 166
QY 475 LNPND-----TSLYDCCINSLOYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520
Db 167 LIENGMIENWIEKLPVCVILTPDCFWGAK---LQGSAYLPGRPDIQMTNLDPEQLLE 223
QY 521 -----HFLYCANAPITFKDGTALALUSCHWADYA 548
Db 224 ELGPFASLEGFRELIDKAQVGOAYVGRPCLDHPPSAPNHHRSQAPNVAHLSGCG 283
QY 549 PVPPF-----LAIGYKGDYSE---AELIMTF-----SLNYPAGDPRLAQAKL 591
Db 284 HGFHKFMWQBELLGGMARPOGELLRAELAQSTFLMSPRLQYEHFRGDIYTHDIGN 343
QY 592 WEEAFLEENRAFQRRMAGMFQVTTAERSLE-DEINRTTAEDLPIPATS-----Y 640
Db 344 SEEQASTVLQAWRRFVQLAQEALPENASQOIHAFFSSTLDDI-LHAFSEVAARVVG 402
QY 641 IVIFLYISALGSYSWSRWVDSKATLGLGVAVVLGAVMAAGFFSYGLTRSSVLILQ 700
Db 403 LMLAYACVTM---LRWD---CAQSQSGVGLAGVLLVALAVASGLGICALLGITFNAATTQ 457
QY 701 VVPELVLSVGADNIFIFVLEYQRLPRRPGEPREHVGIRALGRVAPSMLLCSLSEALCEPL 760
Db 458 VLFFLALGIGVODVFL--LAHAFTEALPGTPIQERMGECQRTGTGSVVLTSINNMAAFILM 515
QY 761 GALTMPAVRTFALTSGLAVILDFLIOMSAFVALLSLSKROEASRLDVCCC----- 812
Db 516 ALVLPALRAFSLOAAIVVGCTFVAVMLVFPAILSLDLRRHRCQLDLVCCFSPSCSAQ 575
QY 813 ---VKPQEL-----PPQOGEGL-- 827
Db 576 VIQILLPQELGDCVTPVGIATLTATVQAFTHCEASSQHVVTILPPQAHLVPPSPDLGSEL 635
QY 828 ---LLG-----PFQKAYAPFLLHMITRGVLLFLA 855
Db 636 FSPGSTRDLLOEBETROKACKSLPCARWNLAHFARYQFAPLQLQSHAKAIVLVFGA 695
QY 856 LFGVLSYMSCHTSVGLDDBALPKDSYLLDYFLNRYFEVGAFFVFTVLGYNFS-SEA 914
Db 696 LIGLSYGATLVODGLATDVVPRGTKEHAFLSAQLRYFSL-YEVALVTQGFDAHQR 754
QY 915 GMAIACSSAGCNPFQTKIQVATFEPGOSYLAIPAS-----SWDDFDIMLTPSSCC--- 967
Db 755 AL-----FDLHORFSSL-----KAVLPPPATQAPRTWLHYRNLQGIQAAFDQ 798
QY 968 ---RLYISGPNKDKCFPSIVNSLNCILKNCMSITMGSVRPSVQFHYLFPWFLND 1018
Db 799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQBPLD-FSOLATTRKLVLD 846
QY 1019 RBNIKCPKGGLAAYSTSVNLTSQGVLASRPMAYHKP----- 1055
Db 847 REGILPPE-LFYMGILTVWSSDPLGLAASQANFYPPPEWLHKDYDTTGEMLRIPPAQ 904
QY 1056 -----LKNSDQYTEALRAARELAANI-TADLRKVPGTDPAPFEPVYTTITNVFY 1102
Db 905 LEPAQFPFLRLGLQXTADPVEAIEGARAACAEAGQVHAYPSGSPF-----LFW 954
QY 1103 EOYLTLPEGLPMLSIC-LVPTFAVSCLLIGLDIRSGLLNLLSIVMLIVTVGPMALWD 1160
Db 955 EOYLGIL--RRCFLAVACILLVCTFLVCALLLNPNWTAGLI-VLVLMAMTVLFGIMGFLG 1011
QY 1161 LSYNAVSLINISVAGMSVEFVSHITRSPASTKPTWLERAKETISMGSAVPAVAMTN 1220
Db 1012 IKLSAIPVILVASGIGVEFTVHVVALGF-LTTOGSRNLRHAHEHTFAPVTDG-AIST 1069
QY 1221 LPGAIVLGLAKAQILQIFPFRLLNLLITLGLHLGLVFLPVLISVYGVPOVNPALALEOKRA 1280
Db 1070 LGLLMLAGSHDFIVRYFFAALTVLTLGLHLGLVFLPVLISILGPP--PEVIQMYKES 1127
QY 1281 BEAANA-----MVASCNHPHSRVSTADNIYN 1308
Db 1128 PEILSPAPQGGGLRWGASSSLPQSFARVTTISMTVAIH 1165

RESULT 30

US-09-293-505-7
; Sequence 7, Application US/09293505
; Patent No. 6348575

GENERAL INFORMATION:

; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 7
; LENGTH: 1182
; TYPE: PRN
; ORGANISM: Mus musculus
US-09-293-505-7

Query Match 8.4%; Score 583.5; DB 2; Length 1182;
Best Local Similarity 24.2%; Pred. No. 2.3e-45;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SPSTHTLLGQ-----PFOG-----WGTWVASWPLTILVLSVIPVVALAAGLVFTLT 376
Db 20 SSAPHILAGSLOAPLWLRAYFOGLLPGLGCRQKHGKVLFLGLVAFGALAGLVRVAVIE 79
QY 377 TDPVELWAPNSQARSEKAFHDQHFQ-PFFRTNQVILTAPNRSYRYSDSLILGPKNFSGI 435
Db 80 TDLEQLAWVEVSRVSGELHYTKELGEBAAVTSQMLI---QTAHQEGGNVLTPE---A 131
QY 436 LDLDLLELLELEQERLHLOWSPQAQNI SLQDI CY--APLNDN-----TSLYDCC 486
Db 132 LD-----LHQAALTASKVQVSLYKGSWDLNKICYKSGVPLIENGMIERIEKLPFCV 184
QY 487 INSLIOYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---FKDGTALA 539
Db 195 ILTPDCFWGAK---LQGSAYLPGRPDIQMTNLDPEQLLEGPASLEGFRELIDKA 241
QY 540 LSCMADYGAPVF-----PFLA---IGGYKGDYS----- 565
Db 242 QVGQAYVGRPCLDPDHPCHPSAPNHRSKQAPNVAQELSGGCHGFSHKFMWQBELLG 301
QY 566 -----BABALIMTF-----SLNYPAGDPRLAQAKLWBEAPLEEMRAFORRMAG 609
Db 302 TARDLQGLLRABEALQSTFLMSPRLQYEHFRGDIYTHDIGNSEEQASNVLOAQWRRFVQ 361
QY 610 MFQVTFTAERSLE-----DEINRTTAEDLPIPATSIV--IFLYISLALGSYSWS 658
Db 362 LAQEALPANASQOIHAFFSSTLDDI LRAFSE-----VSTRVVGVLMLAYACVTMLRWD 417
QY 659 RMVDSKATLGLGVAVVLGAVMAAGFFSYGLTRSSVLILQVVPFLVLSVGADNIFIFV 718
Db 418 --CAQSQAGVGLAGVLLVALAVASGLGICALLGITFNAATTQVLPFLALGIGVDDIFLLA 475
QY 719 LEYQRLPRRPGEPREHVGIRALGRVAPSMLLCSLSEATCFPLGALTMPAVRTFALTSL 778
Db 476 HAFTKAP--PDTPLPERMGECRLSRSTGTSTVALTSVNNMVAFMAALVPIPALRAFSLQAAI 533
QY 779 AVILDPLOMSAFVALLSLSKROEASRLDVCCC-----VKPQE----- 817
Db 534 VVGNFAAVMLVFPAILSLDLRRHRCQLDLVCCFSPSCSAQVIQMLPQELGDRVAVGI 593
QY 818 ---LPP-----PG-----QGEQ-- 826
Db 594 AHLTATVQAFTHCEASSQHVVTILPPQAHLLSPASDPLGSELYSPGSTRDILLSBEGTG 653
QY 827 -----LLLGFFQKAYAPFLLHMITRGVLLFLALFGVLSYMSCHISVGLDQ 873
Db 654 PQAACRPLCAHWTLAHFARYQFAPLQLQSHAKAIVLVFGALLGLSLYGATLVQDGLAL 713


```
Db 977 LVACALLLLSPWTAGLI-VLVAMMTVELFGIMFGIKLSAIPVVLVASIGVFEVTH 1035
QY 1185 ITRSPAISTKPTWLERAKEATISMGSAVPAGV---AMTNLPGLVLGLAKAQLIQIFFR 1241
Db 1036 VALGFLTSHGRNLRRAA-----SALEQTFAPVTDGAVSTLLGLMLAGSNFDFIIRYFPV 1090
QY 1242 LMLLTLLGLLGLVPLVILSVGP 1267
Db 1091 VLVTLTLLGLLGLLPLVLLSLGP 1116

RESULT 32
US-09-248-796A-16160
; Sequence 16160, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16160
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16160

Query Match 8.4%; Score 581.5; DB 2; Length 531;
Best Local Similarity 28.9%; Pred. No. 8.6e-46;
Matches 151; Conservative 100; Mismatches 220; Indels 51; Gaps 16;

QY 642 VIFYLSIALGSSYSSRWVDSKATLGLGGV-----AVVLGAVMAAGPFYSYLG 691
Db 13 IVFGVYVYICF--FSPWQVTKQLVLIGENKIYIRVFVYNNIICNGV---SGIFSII 67
QY 692 IRSSVILVQVPLVLSVADNIFIVLEYQRLPR-RGEPREVHIGALGRVAPSMLLC 750
Db 68 LRSTLIIAIEVIFPLAIGIDNIFIVHELHVISEGNPNLALEVRISQALKHIGSCFIS 127
QY 751 SLSEALCFGLALTPMPAVRTEALTSGLAVILDFLQMSAFVALLSLDKROEASRLDVC 810
Db 128 AVLQVCMFLATSVGMPAVKVFAYYGAGAVLNFSLQMTCFGLLALDQRLDNRVDV 187
QY 811 --CCVKP-----QBLPPPGQGGELLLGFFQKAYAPFLHMTIRGWLILLFLALFGVSIY 862
Db 188 PWTISPIQLQNDDEIDEPVHLEYNFSRWIGDHYAPFLKTKTKPKVITFLVWVGISLS 247
QY 863 SMCHISVGLDQELALPKDSYLLDYFLNRYPEVAGPVYVTTLYNFSSEKGMMAICSS 922
Db 248 LFPKQLGLDQRIATIPSKSYLVNYSVYDVLNVPVFFVVK-DLDYSERLNQKICGG 306
QY 923 -AGCNNEFSQKIQYATEPEOSYLAIASSWVDVDFIDWLT-SSCCRLYISG--PNKD 977
Db 307 FSACDEFSLANTLEQFKRSDISMLSEBASNWLDDFFSGLNLPDLQCCCRPKKSTVFEKTP 366
QY 978 KPCPSTVNSLNLCKNCSITMGSRFSPVEQFKHLYP-----WFLND---RPNIKCPKGG 1029
Db 367 EPCSPNAPQRC-QSCYL---NNHPYDSSMKAFPERDFMYFNDWIOBSPDCPLGGK 421
QY 1030 AAYTSVNLTSQGVLAGRFMAHYHKPKNSQDYTEALRAARELANIATDLRKVPETDPA 1089
Db 422 AAHQQAISRTTE-KIDSSYFRTSPAPLRQDEFIYAYKSGNNIVKEITKLI-----PS 473
QY 1090 FEVRPYTITNVEYOYLITLPGFLWLSICLVPTFAVSCILL 1131
Db 474 MDVFAVSPFFIFFTQYQINVLITVALLTVAMLIYVISTFFL 515
```

RESULT 33

```
US-09-207-857-2
; Sequence 2, Application US/09207857
; Patent No. 6309879
; GENERAL INFORMATION:
; APPLICANT: Buncroft, David A.
; TITLE OF INVENTION: HUMAN PATCHED GENES AND PROTEINS, AND USES RELATED
; TITLE OF INVENTION: THERETO
; FILE REFERENCE: ONV-05001
; CURRENT APPLICATION NUMBER: US/09/207,857
; CURRENT FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: human
US-09-207-857-2

Query Match 8.4%; Score 580; DB 2; Length 1203;
Best Local Similarity 22.9%; Pred. No. 5e-45;
Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;

QY 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLQQFPQG-----WGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSYTPPARTAAAPQLAGSLKAPLWLRAYFQGLLFSLGGCQIRHCGKVLPLG 62
QY 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHG--PFRTNQVILTPNR 417
Db 63 LLAFGALALGRMAIETNLEQLWVEGSRVSQELHYTKELGEEAAVTSQMLIQTAQRE 122
QY 418 SSVYDSLLGKPNFSGILDLLLELELELERHLQVMSPEAORNISLQDICY---AP 474
Db 123 GEN-----ILTPALG-----LHLQAALTASKVQVLSYKSWDLNKKICYKSGVP 166
QY 475 LNPON-----TSLYDCCINSLLQVFNQNRLLLTLLTANQTLMGQTSQVDWKD----- 520
Db 167 LIENGMIERIEKLPCCVILTPDCFWEGAK---LQGSAYILPGRPDILQWNLDPQLLE 223
QY 521 -----HFLYCANAPLTFKDGFTALALSCMADYGA 548
Db 224 ELGPPFASLEGFRELIDKAQVQAVYGRPCLHPDDDLHCPSPAPNHHRSQAPNVHLSGCG 283
QY 549 PVPEP-----LAIGYKGYKDYSE---ASALIMTF-----SLNYPAGDPRLAQA 591
Db 284 HGFSHKFMHQBELLGGMARDPOGELLRAEALQSTFLLMSPRQLYEHPRGDIYQTHDIG 343
QY 592 WEEAFLEEMRAPQRRMAGMFQVTFTAERSLS-DEINRTTAEDLPITFATS-----Y 640
Db 344 SEEQASTVLQAWQRRFVQLAQEALPENASQOIHAFSSITLDDI-LHAFSEVSAARVVG 402
QY 641 IVIFLYISLALGSSYSSRWVDSKATLGLGGVAVVLGAVMAAGPFYSYLGIRSSLVILQ 700
Db 403 LLMLAYACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGCLALLGITTFNAATQ 457
QY 701 VVPFLVLSVGDNIFIVLEYQRLPRRPGEPREHIGRVAIPSMLLCSISEAICFEL 760
Db 458 VLPFLALGIGVDDVFL--LAHAFTALPGTLPQRMWEGECLORTGTSVVLTSINNMAAF 515
QY 761 GALTMPMPAVRTFALTSGILAVILDFLLQMSAFVALLSLDKRQESRLDVCCE----- 812
Db 516 AALVPIPALRAFSLOAAIVVGCTFVANVLPFAIPFSLDRRHHQRDLVLCFSPSSCSAQ 575
QY 813 ---VKPQL-----PPPGQEGEL-- 827
Db 576 VIQILPQELGDTVPVIGIAHLTATVQAFTHCEASSQHVVTILPQOHLVPPSPDPLGSEL 635
QY 828 -----LLG-----PFQKAYAPFLHMTIRGWLILLPLA 855
Db 636 FSPGSGTRDLLGQBEETKQAKCKSLPCARWNLAHFARYQAPAPLLQSHAKAIVLVLFGA 695
QY 856 LFGVLSYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVAGPVYFVTTLYGYNFS-SEA 914
```

Db	696	LLGLSLYGATUVQDGLATDVVPRGTEKHAFLSAQLRYFSL-YEVALVTQGGFDYAHQSQR	754
Qy	915	GMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPAS-----SWDDFDIMLTPSSCC---	967
Db	755	AL-----FDLHQRFSSL-----KAVLPPPATQAPRTWLHYRNWMLQGIQAAPDQ	798
Qy	968	-----RLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRPSVEQHKYLPWFIND	1018
Db	799	DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQEPLD-FSQLTRKLVD	846
Qy	1019	RPNIKCPKGGGLAAVSTSVNLTSKGQVLASRFMAVHKP-----	1055
Db	847	REGLIPPE--LFYMGLTVMVSSDPLGLAASQANFYPPPEWLHDKYDTTGENLRIPPAQP	904
Qy	1056	-----LKNQSDYTEALRAARELAANI-TADLRKVPGTDPAFVEFPPYITINVPY	1102
Db	904	LEFAQFPFLLRGLQKTADFVEAIEGARAACAEAGQAGVHAYPSGSPF-----LFW	954
Qy	1103	BOYLTIPEGULFMLSIC--LYPTFAVSCILLGLDLRSGLLNLLSVMLVDTVGFMAIWD	1160
Db	955	EQYLG--RRCPFLAVCILLVCTFLVCALLLNPMWAGLI-VLVLAMTVELFGIMFLG	1011
Qy	1161	ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1012	IKLSAIPVVILVASGIGVEFTVHVALGF-LTQGSRLRAAAHALEHTFAPVTDG-AIST	1069
Qy	1221	LPGLVLGLAKAQLIQIPFRNLNLLITLLGLHLVFLPVLISYVGPDPNPALEQKRA	1280
Db	1070	LLGLMLAGSHDFIVRYFFAALTVLTLGLLHGLVLLPVLISILGPP--PEVIQMYKES	1127
Qy	1281	EAVAAV-----MVASCNPSPSRVSTADNIYVN	1308
Db	1128	PEILSPAPQGGGLRWGASSSLPQSFAVTTTNTVAIH	1165
RESULT 34			
US-09-909-280A-2			
; Sequence 2, Application US/09909280A			
; Patent No. 6605700			
; GENERAL INFORMATION:			
; APPLICANT: Buncroft, David A.			
; TITLE OF INVENTION: HUMAN PATCHED GENES AND PROTEINS, AND USES RELATED			
; TITLE OF INVENTION: THERETO			
; FILE REFERENCE: CIBT-P02-050			
; CURRENT APPLICATION NUMBER: US/09/909,280A			
; CURRENT FILING DATE: 2001-07-19			
; PRIOR APPLICATION NUMBER: US 09/207,857			
; PRIOR FILING DATE: 1998-12-08			
; PRIOR APPLICATION NUMBER: US 60/067,940			
; PRIOR FILING DATE: 1997-12-08			
; NUMBER OF SEQ ID NOS: 6			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 1203			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-909-280A-2			
Query Match 8.4%; Score 580; DB 2; Length 1203;			
Best Local Similarity 22.9%; Pred. No. 5e-45;			
Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;			
Qy	306	RVPARDKSKMDPKKGTSLSDKL--SFSTHTLLGQFFQG----WGTWASWPLTLVLVS	359
Db	3	RSPPRLPSPSTPPTARTAPQILAGSLKAPLWRAFYQGLPFLSLGCGIQRHCGKVLFLG	62
Qy	360	VIPWVLAAGLVTELTDPVELWSAPNSQARSEKAPHQHG--PFRTNOVILTAPNR	417
Db	63	LLAFGALALGLRMAIETNLEQULVWVGSRVSOELHYTKKGGBEAATYSQMLIQARQE	122
Qy	418	SSRYDSLLGPKNFSGILDLLLELLELLEQLERHLQWMSPEAQRNISLQDICY---AP	474


```
RESULT 35
US-09-807-007-1
; Sequence 1, Application US/09807007
; Patent No. 6881833
; GENERAL INFORMATION:
; APPLICANT: ZAPHIROPOULOS, Peter et al.
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALING PATHWAY
; FILE REFERENCE: 2921-0130P
; CURRENT APPLICATION NUMBER: US/09/807,007
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-007-1

Query Match      8.4%; Score 579; DB 2; Length 1146;
Best Local Similarity 23.3%; Pred. No. 5.7e-45;
Matches 276; Conservative 161; Mismatches 452; Indels 298; Gaps 40;

QY 306 RVAPARDKSMVDPKGTSLSDKL--SFSTHLLGQFFQG---WGTWVASWPLTILVLS 359
DB 3 RSPPLRELPSVTPPARTAAPQILAGSLKAPLWLRAYFQGLLFLSLGCGIQRHCKGVFLG 62
QY 360 VIPVVALAAGLVFTELTPDVELMSAPNSQASEKAFDQHFQ--PFFRTNQVILTAPNR 417
DB 63 LIAFAGALAGLRMAIETNLEQLWVEVGSRSQELHYTKELGEEAAYTSQMLIQTAQOE 122
QY 418 SSRYVDSLLGPKNFSGILDLLELLELQERLRHLQVSPSAQRNLSQDICY---AP 474
DB 123 GEN-----ILTEPALG-----LHLQALTASKVQVSLYKGSWDLNCKYKSGVP 166
QY 475 LMPDN-----TSLVDCCINSLLQVPPNNRTLLLTANQTLMGQTSQVDWKD----- 520
DB 167 LIENGMIERMIKLPFCVILTPDCFWGAK---LQGSAYLPGRPDIQWTNLDPEQLIE 223
QY 521 -----HFLYCANAPLTFKDGITLALSLCHADYCA 548
DB 224 ELGPASLEGFRELIDKAQVQAYVGRPCPLHPDDLHCPPSPAPNHHRSQAPNVABELSGGC 283
QY 549 PVFPF-----LAIGYKGDYSE---AEALIMTF-----SLNNYPAGDPRLAQAKL 591
DB 284 HGFSHKFWHQBELLLGGWARPQOGLLRALALQSTFLMSRPLYEHFRGDIYTHDIGW 343
QY 592 WEEAFLEENRAFQRMAGMFQVTTAERSLE-DEINRTTAEDLPIFATS-----Y 640
DB 344 SEEQASTVLQAMQRRFVQLAQEALPENASQIHFASSTTLDDI-LHAFSEVSAARVVG 402
QY 641 IVIFLYISALGYSWSRWVDSKATIGLGAVVVLGNVAMGFFSVLGRSSLVILQ 700
DB 403 LMLAYACVTM---LRWD---CAQSGSVGLAGVLLVALAVASGLGALLGITFNAAITQ 457
QY 701 VVPFLVLSGADNIFIFVLEYQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCPFL 760
DB 458 VLFPALGIGVDDVFL---LAHAFTEALPGTPOERMGECLOQTGTSVLTSLNNAAFLM 515
QY 761 GALTMPAVRTPALTSGLAVIDLFLQMSAFVALLSLDSKROEASRLDVCCC----- 812
DB 516 AALVFPALRAFSLQAAIVVGCTFFVAVMLVFPAILSLDLRRHRCQRLDLVLCFSSPCSQA 575
QY 813 ---VFPQEL-----PPGQGEGL-- 827
DB 576 VIQILPQELGDTVPVGIHAHLTATVQAFTHCEASSQHVVTILPQAHVLPVPSDPLGSEL 635
QY 828 -----LLG-----PFQKAYAPELLHMITRQGVLLFLA 855
DB 636 FSPGSTRDLQGEETROKACKSLPCARWNLAHFARYQFAPLQLLQSHAKAIVLPLFGA 695
QY 856 LFGVLSYSNCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTTLGYNFS-SEA 914
```

```
DB 696 LLGLSLYGATLVQDGLALTDVVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDYAHQSQR 754
QY 915 GMAICSSAGCNFSPFTQKIQIATEPEQSYLAIPAS-----SWDDFDLWLTSSCC--- 967
DB 755 AL-----FDLHORPSSL-----KAVLPFPATQAPRTWLHYRNMLOGIOAAFDQ 798
QY 968 -----RLYISGNKDKFCFSTVNSLCKNCKMSITWGSVRPSVEQPHKYLWPLND 1018
DB 799 DWASGRITRHSYRG-SED-----GALAYKLLIQTGDAQELLD-FSOLTTRKLVLD 846
QY 1019 RPNIKCPKGLAAYSTSVNLTSDGQVLASRFWAYHKP----- 1055
DB 847 REGLIPPE--LFYMGILTVMVSSDPLGLAASQANFYPPPEPEMLHDKYDTTGENFRIPPAQP 904
QY 1056 -----LKNSQDYTEALRAARELAANI-TADLRKVPCTDPAFVFPFTTINVFY 1102
DB 905 LEFAQPPFLRLGLQKTADFEAIEGARACACABAGQAGVHAYPSGSPF-----LPW 954
QY 1103 EGYLTILPEGLFMLSIC--LVPTFAVSCILLGLRLSLNLLLSIVMLVDTVGFMAWLD 1160
DB 955 EGYLGL--RRCFLAVCILLVCTFLVCALLLNPNWTAGLI-VVLAMTVFELFGMGFLG 1011
QY 1161 ISYNAVSLINLVSAYGMSVFVSHITRSPFAISKPTMLERAKEATISMGSAVPAVMTN 1220
DB 1012 IKLSAIPVVILVASVIGVETVHVVALGF-LTTQGSRLRAAAHLEHTFAPVTDG-AIST 1069
QY 1221 LRGILVLGLAKAQLIQIPFRNLNLLITLILGLHGLVFLPVILSYGVP 1267
DB 1070 LIGLLMLAGSHPDFIVRYFFAALTVLTLGLLHGLVLLPVLISILGP 1116

RESULT 36
US-08-540-406-4
; Sequence 4, Application US/08540406
; Patent No. 5837538
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-540-406-4
```


Db 456 IRSQAGVGIAGVLLSITVAAGLGFALLGIPFNASSQIIVPFLALGLGVQDMFLLTHTY 515
Qy 722 QRLPRRPGEPREVIHGRALGRVAPSMILCSSEACFFFLGALTMPAVRTFALTSLGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGLSVLLASLCNVMFLAAALLPAPRPFVCLQAAIILL 573
Qy 782 LDFLLQMSAFVALLSDSKRQASRLDVCCCKQELPP----- 820
Db 574 FNLGSILLVFPAMISLDLRRSAAADLLCCLMPESPUPKKKIPERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGGL-----LLGFFQKAYAPFLLHMITRGVLLLFALFGVSLYSMC 865
Db 634 TRQPLDPDVSENVTKTCLSVLSLTKWAKNQAPFIMRPAVKVTSMLIAVILTSVMGAT 693
Qy 866 HISVGLDQELALPKDSYLLDVPFLNR---YFVGAPVYFVTTLGYFNSSAGNAICSS 922
Db 694 KVKGLDLTDIVPENT--DEHEFLSRQEKYGF-----YN-----MYAVTQ- 732
Qy 923 AGCNNSFP--TKIQIYATEFPEQSYLAIP-----ASSWVDDFIDML----- 961
Db 733 ---GNFEYPTNQKLLY--EYHQD--FVRIPNIIKNDNGGLTKFWLSLFRDMLLDLQVAPDK 786
Qy 962 -TPSSCCRLYISGPNKDFPCPSTVNSLNCMLKMSITMGSVRPSVEQ----- 1007
Db 787 EVASGCI-----TQEYWCNKASDE-GILAYKLVQVGHVDPNDKSLITAGHRLVDK 838
Qy 1008 -----FHXYL-PWFLNDRPNIKCPKGGGL-----AYSTSVNLTSDQVULASRPM 1050
Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPQORWIHSPBDVHLEIKKSSPLIYTQLP 898
Qy 1051 AYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEPVYTTNVFVEQVLTILP 1110
Db 899 FYLSGLSDYKSTIKTILSRVDRDLCCKVEA--KGLPN-----FPGSIFLFWQVLYLRT 949
Qy 1111 EGLFMLSCLVPTFAVSCILLGLDRSGLNLLSIVMILVDTVGFMALMDISYNVSLIN 1170
Db 950 SILLALACALAAVP-IANVLLLNAAVAVLTALATLVLLGVMALLGVKLSAMPAVL 1008
Qy 1171 LVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAFVAGVAMTNLPGLILVGLA 1230
Db 1009 LVLAIGRVHFTVHLCLGVTSTGC-----KRRASLALES-VLAPV---VHGALAAALA 1059
Qy 1231 KAQL-----IQIFFRNLALTLLGLAHGLVFLPVLVSYPG--DVNPALALEQKRAE 1281
Db 1060 ASMLAASECGFVARLFLRLDOLVFLGLDGLLFFPVLVSILGPAAEVTRI----- 1110
Qy 1282 EAVAAVMVASCPNHPSPRST 1301
Db 1111 -----EHPERLST 1118

RESULT 39
US-08-918-658-4
; Sequence 4, Application US/08918658
; Patent No. 6429354
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; GOODRICH, LISA V
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,658
FILING DATE: 22-Aug-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/656,055
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/540,406
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-918-658-4
Query Match 7.6%; Score 526.5; DB 2; Length 1311;
Best Local Similarity 22.2%; Pred No. 7.1e-40;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;
Qy 336 ILGQFFQGWGTWASWPLTILVLSVIPVVALAAGLVFTLTTPVELMSAPNSQARSEKA 395
Db 72 ILGFLQDGACKVLF--VAILVSTFCV-----GLKSAQIHTRVDQLWVQGGRLAEALK 124
Qy 396 FHDQHFPP-FRTNQVILTAPNRSYRSDLSLLGPNFSGILDLLELLELLELRLHL 454
Db 125 YTAQALGEADSSHTQLVI---QTAQKPDVSLHP---GAL-----LEHL 162
Qy 455 QVMSPEAQRN-----SLQDICVAPLNPD-----NTSLVD-----CCINSLLQYFQNN 497
Db 163 KVVHAATRTVMYDIEWRLKDLCSYSPIDPFGVHHIESIDNVIPCAITPLDCFWEG 222
Qy 498 RTLL-----LTDNQTLMGQTSQ----- 515
Db 223 SKLLGPDYPIYVPHLKHKLQWHLNPLEWBEVKKLKFQFPLSTTIEAYMKRAGITSAYMK 282
Qy 516 ---VDWKDHFLYC--ANAPLTFKDTALALSCWADYGAPVFP-----LAIGG----- 558
Db 283 KPCLDFTDP--HCPATAP-NKKSGHIPDVAABELSHGCGYFAAAYMHWPEQLIVGGATRNS 339
Qy 559 ---YKGDYSEAEALIMTFSLLNYPAGDPRLAQAKLWEEAFLEEMRAFQREMAGMFQVT 614
Db 340 TSALKRARKLQTVVQLMGEREMYEYWDHYKHVQIOWNQEKAAAVLDQWQKFAAEVRKI 399
Qy 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSWRVM 661
Db 400 TTSGSVSSAYSFPYSTSTLNDILGKFSEVSLKNIILGYMFMFLIYVAVTL---IQW-RDP 455
Qy 662 VDSKATILGCVAVVLGVAWAMGFFSYLGRSSVLVQLVVPFLVLSVCADNIFIPVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLGFALLGIPFNASSQIIVPFLALGLGVQDMFLLTHTY 515
Qy 722 QRLPRRPGEPREVIHGRALGRVAPSMILCSSEACFFFLGALTMPAVRTFALTSLGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGLSVLLASLCNVMFLAAALLPAPRPFVCLQAAIILL 573
Qy 782 LDFLLQMSAFVALLSDSKRQASRLDVCCCKQELPP----- 820
Db 574 FNLGSILLVFPAMISLDLRRSAAADLLCCLMPESPUPKKKIPERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGGL-----LLGFFQKAYAPFLLHMITRGVLLLFALFGVSLYSMC 865
Db 634 TRQPLDPDVSENVTKTCLSVLSLTKWAKNQAPFIMRPAVKVTSMLIAVILTSVMGAT 693

QY 1282 EAVAVMVASCNHPHRSVST 1301
DB 1111 -----EHPERLST 1118

RESULT 42

US-09-754-032-4
; Sequence 4, Application US/09754032
; Patent No. 6921646
; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P
; JOHNSON, RONALD L
; ADDRESS: FOUR EMBARCADERO CENTER, SUITE 3400
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: US
; ZIP: 94111

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,032
; FILING DATE: 03-Jan-2001

CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995

ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-754-032-4

Query Match 7.6%; Score 526.5; DB 2; Length 1311;
Best Local Similarity 22.2%; Pred. No. 7.1e-40;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 LIGQFQGGTGWASWPLTILVLSVPPVLAAGLVFTLTDPVELWSPNSQARSEKA 395
DB 72 ILGCLQGDAGKVLFP--VALLVLTFCV-----GLKSAQHTRVDQWVQEGGRLEALK 124
QY 396 FHQHFQGP-FTNQTIVLTAPNRSYRYSLLGPNFSGILDLDLLELLELRLRL 454
DB 125 YTAQALGADSSHTQLVI---QTKADPDVSLHF---GAL-----LEHL 162
QY 455 QWSPSEAGQNI-----SLQDICVAPLNPDP-----NTSLYD-----CCINSLLQYQNN 497
DB 163 KVVHAATRVTHMYDIEWRLKDCYSPSPDPPEGYHHIESIIDNVIPCAIITPLDCFWEG 222
QY 498 RTLL-----LITANQTLWGQTSQ--- 515
DB 223 SKLLGPDPIYVPHLKHKLQWHLNPLNPLEVVEVKLKFQFPLSTTEATWKGAGITSYMK 282
QY 516 ---VDMKHFLYC-ANAPLTFKDGFTALALSCWADYGAPVFPF-----LAIGG----- 558

DB 283 KPCLDPTDP--HCPATAP-NKSGHIPPVAAELSHGCGYFAAAAMHMPQLIVGATRNS 339
QY 559 ---YKGDYSEAEALIMTFSLNNYPAGDPPLAQAKLWEEAFLEEMRAFORWAGMCFQVT 614
DB 340 TSALKRARKLQTVVQJGGEREMETWADHYKHVQHGWNOEKAAAVLDDAWORFAAEVRKI 399
QY 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSSRVM 661
DB 400 TTSQSVSSAYSPPYFSTSTLNDILKFSRVSLSKNILGYMFMILYVAVTL---IQW-RDP 455
QY 662 VDSKATLGLGCVAVVGLGAVMAAGPFSYLGIRSSILVILQVVPFLVSVGADNIFIVLSY 721
DB 456 IRSQAGVGIAGVLLSITVAAGLGFALIGIPFNASTQIVPFLALGLGVQDMFLITHTY 515
QY 722 QRLPRPGEPREHVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPAVRTEALTSGLAVI 781
DB 516 --VEQAGDVPREERTGLVKKSGLSVLLASLCNVWAFALAAALPIPAFRVFCLOAAILL 573
QY 782 LDFLLQMSAFVALLSLDSKQBEASRLDVCCVCKPQELPP----- 820
DB 574 FNLGSILLVFPAMISLDLRRSARADLLCCLMPESPPLPKKI PERAKTRKNDKTHRIDT 633
QY 821 -----PCQGEGL-----LLGFFQKAYAPFLHMITRGVVLILLFALFGVSIYSMC 865
DB 634 TRQPLDPDVSENVTKTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTSWGTAT 693
QY 866 HISVGLDQELALPKDSYLLDYFLPLNR---YFVGAPVVFVTLGLVNFSSBAGMAICSS 922
DB 694 KYDGLDLDTDIVPENT---DEHEFLSRQEKYFGP-----YN-----MYAVTO- 732
QY 923 AGCNMFSE--TKIQYATEPPEQSYLAIP-----ASSWVDDFDIDL- 961
DB 733 ---GNFEYPTNQKLY--EYHDQ-FVRIPNIIKNDNGGLTKFMLSIFRDLMLDLQVAPDK 786
QY 962 -TPSSCCRLYISGPNKDKFCPSTVNSLNCNMCNMCITMGSVRPSVEQ----- 1007
DB 787 EVASGCI-----TQEWCKNASDE-GILAYKLMVQTHVDNPDKSLITAGHRLVDKD 838
QY 1008 -----FHKYL-PMFLNDRPNIKCPKGL-----AAVSTSVNLTSDQVLSRPM 1050
DB 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPQORWIHSPEDVHLEIKSSFLIYTQLP 898
QY 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFYEQLTILP 1110
DB 899 FYLSGLSDTXXSIKTLIRSVDLCLEYA--KGLPN-----FPGSIPFLFWEQIYLRT 949
QY 1111 BGLFMLSICLVPTFAVSCILLGLDLSRGLNLLSIVMLVDTVGFPMALWDISYNAVSLIN 1170
DB 950 SILLALACALAAVF-IAVMVLLNAAVAVLTALATLVQLLGYMALLGVKLSAMPAVL 1008
QY 1171 LVSAVGMSEVESHITRSPALSTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLA 1230
DB 1009 LVLAIGRGVHPTVHLCLGFTVSIGC-----KRRASLALES-VLAPV-----VHGALAAALA 1059
QY 1231 KAQL-----IQTPFFRLNLLITLLGLLHGLVFLPVLISYVGP--DYNPALALRQKRAE 1281
DB 1060 ASMLAASECGFVARLFLRLLDIVFLGIDGLLFFPIVLSILGPAAEVRPI----- 1110
QY 1282 EAVAVMVASCNHPHRSVST 1301
DB 1111 -----EHPERLST 1118

RESULT 43

US-08-916-140-4
; Sequence 4, Application US/08916140
; Patent No. 6946257
; GENERAL INFORMATION:
; APPLICANT: Scott Mathew P.
; APPLICANT: Goodrich, Lisa V.
; APPLICANT: Johnson, Ronald L.
; APPLICANT: Epstein, Ervin Jr.

[illegible]

667 FQHYTFELMRWVKFLTWGFLAAALISSYASTRLQDGLDIIIDLVKDSNEHKFLDAQTR 726
893 YFEGAPVYFVTTGLGYNFSEAGMAICSSAGCANNFSTQKIQYATEPEQSYLAIP--- 949
727 LF--GFYSMYAVTQG-----NFEYPTQQLLRDY-HDSFVRVPHVI 764
950 -----ASSWDDFIDMLTPSSCCRLYISGPNKDKFCFSTVNSLNCNKC----- 993
765 KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEBYRDGRLTKECWFPNASSDA 812
994 -----MSITMGSRVRSVEQ-----PHKYL-PWFLNDRNPKCPKGG 1028
813 ILAYKLIQVOTGHVNDPNVDKELVLTNRLVNSDGIINORAFYNLSAW-----ATNASSPTL 868
1029 LAA--YSTSVNLTSGQVLASRFMAHYHKPLKNSQD-----YTEALRAAR 1070
869 LEANCIRNANGASQGLKPYERQYFQHP--NEYDLKTPKSLPYVAQMPFYLHGLTDS 926
1071 ELAANITADLRKVPQTDPAFV--PPYITITNVEYQYLTILPEGLFMLSCLVPTFAVSC 1128
927 QI-KTLIGHIRDLVKYEGFGLPNYPGIPFIFWEQYMT-LRSSLAMILLACVLLAALV 984
1129 LLLGLDLRGLNLLISVILVDTVGFMAWDISTNVLNLSVAGMSVEFVSHITRS 1188
985 SLLLSVMAAVILSVLASLAQIFGAMTLGILKLSAIPAVILILSVGNMLCF--NVLIS 1042
1189 PAISTKPTWLERAKATTSMGSAVAGVAMTNLPGILVLGLAKAQIQIFPRLNLLITL 1248
1043 LGFMTSVGNRRVQSLQMSGLPIVHGMLTSGVAVFMLSTSPPEFVTRHFCWLLLVLC 1102
1249 LGLLHGLVPLPVLSYVGVDPVNPALALEOKRABEAAVAVMVASCPNHPSTVSTADNIYN 1308
1103 VGACNSLLVFPILLSMVGEAE-LVPLE-----HPRISTPSPLPVR 1143
1309 HS 1310
1144 SS 1145

RESULT 47

US-08-674-509B-48

Sequence 48, Application US/08674509B

Patent No. 6261786

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tabin, Clifford J.

APPLICANT: Marigo, Valeria

TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEGOG AGONISTS

TITLE OF INVENTION: AND ANTAGONISTS

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESSES:

ADDRESS: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674,509B

FILING DATE: 02-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/460,900

FILING DATE: 05-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.06
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1299 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-674-509B-48

Query Match 6.7%; Score 464.5; DB 2; Length 1299;
Best Local Similarity 20.4%; Pred. No. 5.4e-34;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

QY 321 KGTSLSDKLSFSTHTLLQPFQCGQGTWVASWPLTILVLSVIPVVALAGLVTELTDTDPV 380
DB 43 KGKAGRSRTAIYLSVFSQSHLETLGSSVQKAGKVLFAILLVLSTFCVGLKSAQIHSKVH 102
QY 381 ELSAPNSQARSEKAFHQHFG-PPFRTNOVILTA---PNBSYRYDLSILLGPKPSGL 436
DB 103 QLWQEGGGLAEALAYTKTIGEDSATHLLIQTTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLEQLERLRLHQLVNSPEAQRNISLQDICYAPLNPDPNTSLY-----DCC 486
DB 154 HLEVIVKATAKVHLYDTE-W-----GLRDMCMESTPSFEGIYIEQLRLHLIPCS 204
QY 487 INSLIQYQN-----NRTLLLTAN-----QTLMGQTSQVDMKDH 521
DB 205 IITPLDCFWEGSQLLPESAVVIPGLNQLRLWTLNLPASVMQYMKQKMESEKISDFETV 264
QY 522 FLYCANAPLTQDGTALALSCM-----ADYCAPVFPFLAIG--GVK 560
DB 265 EOYMKRAAIA---SGYMEKPCLNPLNPNCPDAPNKNSTQPPDVG-----ILSGCYGYA 317
QY 561 GKDYSEAEALIMTFLSNYPAGDPLAQ-----589
DB 318 AKHMHWPPEELIVGGAKRN-RSGHLRKAQALQSVVQLMTEKEMYDQWQNYKVHHLGWQ 376
QY 590 -----KLWEEAP---LEEMRAPQRMAGMFQVTTAERSLEDEINRTTAEDLPATSY 640
DB 377 KAAEVLNAWQNFREVEQLLRKQSRITATNDIYVFSAAALDDILAKFSHPALSIVIGV 436
QY 641 IVIELYISLALGSSYSSRWVDSKATLGLGVAVVLGAVMAAMGFPFVGLTSSILVILQ 700
DB 437 AVTVLY---AFCTLLRW-RDPVQGSSVGVAGVLLMCFSTAGLGLSALLGIVFNAASTQ 492
QY 701 VVPLVLVSGADNIPFVLEYOQLPRRPGEPREHIGRALGRVAPSMELCSLSEACFPFL 760
DB 493 VVPLALGLGVDFHIFMLTAAYAESNR--EQTCL-----ILKKVGPSSILFSACSTAGS 546
QY 761 GALTMPMPAVRTFALTSLGLAVILDFLLOMSAFVALLSDSKRQESRLDV-CCCV-----K 814
DB 547 AAFIPVPALKVFCQAIAIVMCSNLAALLVFPAMISLDRRTTAGRADIFCCCPVWKEQ 606
QY 815 PQELPP-----PQOEGEGL-----LLGFF 832
DB 607 PKVAPPVPLPNNNGRGARHPKSCNNNRVALPAQNPLEQRADIPGSSHSLASFSLATFA 666
QY 833 QKAYAPFLHWTIRGVLLLFALFGVSLYSMSCHISVGLDQELALPKDSYLLDYLFLNLR 892
DB 667 FQHTYPLMRSWRFLTWGFLAALISLYASTRIQDGLDIIIDLVKDSNEHKFLDAQTR 726
QY 893 YFEGAPVYFVTTGLGYNFSSBAGMAICSSAGCANNFSTQKIQYATEPEQSYLAIP--- 949
DB 727 LF--GFYSMYAVTQG-----NFEYPTQQLLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWDDFIDMLTPSSCCRLYISGPNKDKFCFSTVNSLNCNKC----- 993
DB 765 KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEBYRDGRLTKECWFPNASSDA 812

994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKG 1028
813 ILAYKLIVQTGHVDNPDVKELVLNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTL 868
1029 LAA--YSTSVNLTDGGOVLASRFMYHKLKNSQD-----YTBALRAAR 1070
869 LRANCIRNRANGASQGLPYPRQYFHP--NEYDLKPKSLPLVYAQMPPYLHGLTDS 926
1071 ELAANITADLRKVPCTDAFV--PPYITNVFVEQVLTILPEGLFMLSICLVPTFAVSC 1128
927 QI-KTLIGHIRDLSVKYEGFGLNVPYSGIPFWEQYMT-LRSSLAMILACVLLAALVLV 984
1129 LLLGLDRLSGLNLLSIYMLVDTVGFMAWDISYNAVSLNLVSAVMSVEFVSHITRS 1188
985 SLLLSVAAVLVLSLASAQIFGAMTLGKLSAIPAVILISVGMMLCF--NVLIS 1042
1189 FAISTKPTMLERAKETISMSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLTTL 1248
1043 LGFTSVGNRRRVQLSQMSLGLPLVHGMLTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102
1249 LGLLHGLVPLVILSYGPDVNPALALQKRAEAAVAAVMVASCNPHSPRVSTADNIYV 1308
1103 VGACNSLLVFPILLSMVGPEAE-LVPLE-----HPDRISTSPPLVR 1143
1309 HS 1310
1144 SS 1145
RESULT 48
US-08-954-698-48
Sequence 48, Application US/08954698
Patent No. 6271363
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,698
FILING DATE: 20-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,386
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 04-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1299 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-954-698-48
Query Match 6.7%; Score 464.5; DB 2; Length 1299;
Best Local Similarity 20.4%; Pred. No. 5.4e-34;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSFSTHTLLGQFFQGWGTVWASWPTILVLSVIPVVAALAGLVFTLTPDPV 380
DB 43 KGKARGSRTAIYRSVQSHLETIGSSVQKHAGKVLFAVLVLTSTFCVGLKSAQIHSHKVH 102
QY 381 ELMSAPNSQARSEKAFHQHFG-PFFRTNQVILTA-----PNRSSYRYDSLLGLPKNFSGIL 436
DB 103 QLMIQEGGGLEAEALAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLEQLERLHLQVMSPEAQRNLSLQDIQYAPLNPNTSLV-----DCC 486
DB 154 HLEVLRKATAVKVHLXYDTE-W-----GLRDMCNMPSFSEGIYYIEQILRHLPCS 204
QY 487 INSLLOVQFON-----NRTLILLTAN-----QTLMGQTSQVDMKDH 521
DB 205 IITFLDCFWEGSQLLGPESAVVIGLNRQLLWTLTPASVMQKQMSBEKISFDPETV 264
QY 522 FLYCANAPLTPKDGTAALALSCM-----ADYGAPVFPPLAIG--GYK 560
DB 265 EQYMKRAAIA--SGYMEKPCLNPLNPNCPDAPNKNSTOPPVG-----ILSGCYGYA 317
QY 561 GKDYSEBALIMTSLNNYPAGDPLAQA-----589
DB 318 AKHMHPEELIVGAKGN-RSGHLRKAQALQSVVQLMTEKEMYQWQNDYKVHGLWTOE 376
QY 590 -----KLWEEAF--LEEMRAFQRMAGMQVFTTAERSLEDEINRTABDLPIFATSY 640
DB 377 KAAEVLANWQNRFSREVEQLLRKQSRATNYDIVVFSAAALDDILAKSPHSALSIVIGV 436
QY 641 IVIFLYISLALGSYSSWSRVNDSKATLGLGGAVALGAVNAAMGFFSYIGRSLVILQ 700
DB 437 AVTVLY--AFCTLLRW-RDPVRGQSVGVAGVLLMCFSTAAGLSALLGVFVNAATQ 492
QY 701 VVPLVLSVGADNIFIVLEYQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAIQFPL 760
DB 493 VVPLFALGLVDHIFMLTAAYAESNRR--EQTKL-----ILKVGPSILFSACSTAGSFFA 546
QY 761 GALTMPAVRTFALTSGLAVIDFLLOMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
DB 547 AAFIPVPALKVFCLOAAIVMCSNLAAALLVFPAMISLDRRTAGRADIFCCCFPPVWKEQ 606
QY 815 POELPP-----PQCEGL-----LLGFF 832
DB 607 PKVAPPVPLPLANNNGRGARPKSCNNNRVALPAQNPLEQRADIPGSSHSLSASFSLATFA 666
QY 833 QKAYAPFLLHMTTRGVVLLLLFLALFGVSLYSNMCHISVGLDQELALPKDSYLLDYFLFLNR 892
DB 667 FQHYTFPLMRSWVKFLTVMGFLAALISLIYASTRIQDGLDIIIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVTTTLGYNFSSEAGNNAICSSACNNNFSFTQKIQYATEPPEQSYLAIP----- 949
DB 727 LF--GFYSMYAVTQG-----NFEVPTQQLLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWVDDFDIMLTPSSCCRLYISGNKDKFCFSTVNSLNLKNC-----993
DB 765 KNDNGGLPDFMLLFSWL-----GNLQKIFDEEYDRDGLTKCECFPPNASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKG 1028
DB 813 ILAYKLIVQTGHVDNPDVKELVLNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTL 868

Db 927 QI-KTLIGHIRDSVKYEGFGLPNYSGIPFIWEQYMT-LRSSLAMILACVLLAALVLV 984
Qy 1129 LLLGLDLSGLNLLSIYMLVDTVGFMAIWDISYNAVSLNLSVAVGMSVEFYSHITRS 1188
Db 985 SLLLSVWAAVLVILSVLASIAQIFGATLLGIKLSAIPAVILSVGMMLCF--NVLLIS 1042
Qy 1189 FAISTKPTWLERAKATISMSGSAVAGVAMTNLPGILVLGLAKAQIIGIFFRNLMLTLL 1248
Db 1043 LGFTSVGNRRQRVQLSQMSLGLPLVHGMILSGVAVFMLSISPFVFIRHFCWLLLVLC 1102
Qy 1249 LGLLHGLVPLVILSYGPDVNPALAEQKRAEEAAVAVMVASCPNPSRSTADNIYVN 1308
Db 1103 VGACNSLLVFPILLSWVGPEAB-LVPLE-----HDIRISTSPLPVR 1143
Qy 1309 HS 1310
Db 1144 SS 1145

RESULT 50

US-09-448-188-48
; Sequence 48, Application US/09448188
; Patent No. 6607913

GENERAL INFORMATION:

; APPLICANT: Ingham, Phillip W.
; McMahon, Andrew P.
; Tabin, Clifford J.

; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; Proteins and Uses Related Thereto

; NUMBER OF SEQUENCES: 48

; CORRESPONDENCE ADDRESS:

; ADDRESSER: FOLEY, HOAG & ELIOT LLP

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109-2170

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/448,188

; FILING DATE: 23-No. 6607913-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/462,386

; FILING DATE: 05-JUN-1995

; APPLICATION NUMBER: US 08/435,093

; FILING DATE: 04-MAY-1995

; APPLICATION NUMBER: US 08/356,060

; FILING DATE: 14-DEC-1994

; APPLICATION NUMBER: US 08/176,427

; FILING DATE: 30-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: HMV-006.12

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-832-1000

; TELEFAX: 617-832-7000

; INFORMATION FOR SEQ ID NO: 48:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1299 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-09-448-188-48

Query Match

Best Local Similarity 6.7%; Score 464.5; DB 2; Length 1299;

Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

Qy 321 KGTSLSDKLSFSTHTLLGQFFQGMGTWVASWPLTILVLSVIPVVALAAGLVFTBLTDPV 380
Db 43 KGRAGSRRTAIYLSVFSQSHLETGSSVQKHAGKVLFVAILVLSTFCVGLKSAQIHSKVH 102
Qy 381 ELMSAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSYRYDSLLGLGPKNFSIL 436
Db 103 QLWIOEGGLEAEALAYTKTIGEDESATHQLLIQTTHDPNASVLHPQALL-----A 153
Qy 437 DLIDLLELLELOERLHLQVMSPEAQRNLSLQDIACYAPLNPDPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDCMNPSTPSFEGIYIYIQLRHLIPCS 204
Qy 487 INSLLOVQFON-----NRTLLILLTAN-----QTLMGQTSQVMDKH 521
Db 205 IITPLDCFWEGSQLLGPESAVVIFGLNQRLLTTLNPASVMQMKQKMSBEKISDPETV 264
Qy 522 FLYCANAPLTFKOGTALALSCM-----ADYGAPVFPFLATG--GYK 560
Db 265 EQYMKRAAIA---SGYMEKPCLNPLNPNCPDTAPNKNSTQPPDVGA---ILSGCYGYA 317
Qy 561 GKDYSEAEALIMTSLANNYPAGDPLAQA-----589
Db 318 AKRMHWPPEELIVGGAQRN-RSGHLRKAQALQSVVQLMTEKEMYQWQNDYKVVHGLWTQE 376
Qy 590 -----KLWERAF---LEEMRAFQRRMAGMPQVTTAERSLEDEINRTTAEPLIFATSY 640
Db 377 KAAEVLANAQWRNFSREVEQLLRKQSRATNYDIYVFSAAALDDILAKSPHSALSIVIGV 436
Qy 641 IVIFLYISLALGYSYSSWSRVWDSKATLGLGCVAVLVGAVMAAGFFSYLGRSLVLILQ 700
Db 437 AVTVLY---AFCTLLRW-RDFVRGQSGVGAVLLMCFSTAGLGLSALLGIVFVNAATQ 492
Qy 701 VVPELVLSVGDADNIFIVLEYORLPRRPGEPREHVHIGRALGRVAPSMCLLSLSEACFPFL 760
Db 493 VVPELALGLGDVHIFMLTAAYAESNR--EQTKL-----ILKKVPSILFSACSTAGSPPA 546
Qy 761 GALTPMPAVRTFALTSLGLAVILDFLQMSAFVALLSLDSKQESRLDV-CCCV-----K 814
Db 547 AAFIPVPALKVFCQAIVMCSNLAAALLVFPAMISLDLRRRTAGRADIFCCCPFPVWKEQ 606
Qy 815 POELPP-----PGQEGEGL-----LLGFF 832
Db 607 PKVAPPVPLPLNNNGRGARHPKSCNNNRVALPAQNPLEQRADIPGSSHSLSASFSLATFA 666
Qy 833 QKAYAPFLLHWITRGVLLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNR 892
Db 667 FQHYTPFLMRSWVKFLTWGFLAALISLYASTRLQDGLDIDLVPKDSNEHKFLDAQTR 726
Qy 893 YFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCANNFSFTQKIQYATEPEQSYLAIP--- 949
Db 727 LF--GFYSMYAVTQG-----NFEYPTQQLLRDY-HDSFVRVPHVI 764
Qy 950 -----ASSWVDDFDIWLTPSSCCRLYISGNKDKFCFSTVNSLNLCKNKC----- 993
Db 765 KNDNGGLPDPFMLLLFSEWL-----GNLQKIFDEEYRDRGLTKECMFPNASSDA 812
Qy 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db 813 ILAYKLIYQTHGVDPNPVKELVLTNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTEL 868
Qy 1029 LAA--YSTSVNLTSQGVQLASRFMAYHKPKNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQGLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQMPFYHLGLTDT 926
Qy 1071 ELAANITADLRKVPCTDPAFEV--PPYTTITNVFYQYLTILPEGLFMLSILCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDSVKYEGFGLPNYSGIPFIWEQYMT-LRSSLAMILACVLLAALVLV 984
Qy 1129 LLLGLDLSGLNLLSIYMLVDTVGFMAIWDISYNAVSLNLSVAVGMSVEFYSHITRS 1188
Db 985 SLLLSVWAAVLVILSVLASIAQIFGATLLGIKLSAIPAVILSVGMMLCF--NVLLIS 1042

QY 1189 FAISTKPTWLERAKEATISMGSAVAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLLITL 1248
DB 1043 LGFMTSVGNRQRVOLSMQMSLGLPLVHGLMTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102
QY 1249 LGLHLGLVLPVLTSVYVGDVNPALALEQKRAEEAAVAVMVASCPNHPBSRVSTADNIYN 1308
DB 1103 VGNCSLLVFPILLSMVGPEAB-LVPLE-----HPDRISTPSPUPVR 1143
QY 1309 HS 1310
DB 1144 SS 1145

RESULT 51
US-08-954-128-48
; Sequence 48, Application US/08954128
; Patent No. 6610656
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,128
; FILING DATE: 20-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/462,386
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 04-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMV-006.12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-128-48

Query Match 6.7%; Score 464.5; DB 2; Length 1299;
Best Local Similarity 20.4%; Pred. No. 5.4e-34;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSFSTHLLGQFGQGWVWASWPLTLLVLVSVIPVVALAGLVTELTDDVP 380
DB 43 KGKARGSRTAIYLRVSFQSHLETGLGSSVQKHAGKVLVFAILLVSTFCVGLKSAQIHSKVH 102

QY 381 ELMSAPNSOARSEKAFHQHFG-PFPRTNQVILTA---PNRSSRYRDSLLGLPKNFSGIL 436
DB 103 QLWIQBGGGLEAEIATVQKTIGEDESATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELELEQLERLHLQVWSPQARNISLQDICYAPLNPDNSTLY-----DCC 486
DB 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCNMPSTPSPEGIYYIEQILRHLIPCS 204
QY 487 INSLLOQYFON-----NRTLALLLTAN-----OTLMQOTSQVDWKDH 521
DB 205 IITPLDCFWEQSOLLGPESAVVIPGLNQRLLTWTLNPASVMQMKMKMSEKISFDFTV 264
QY 522 FLYCANAPLTPKDGTAALASC-----ADYGAPVFPFLAIG--GYK 560
DB 265 EQYMKRAAIA---SGYMEKPCLNPLNCPDTPAPKNKSTQPPDVGA---ILSGCGCYGA 317
QY 561 GKDYSEAEALIMTFPSLNNYPAGDPRLAQA----- 589
DB 318 AKHMHMPEELIYGGAKRN-RSGHLRKAQALQSUVQLMTEKEMYDQMDNYKVHHLGWTOE 376
QY 590 -----KLWEAP---LEEMRAFORRMAGMFOVPTTAERSLEDEINRTTAEDLPIATSY 640
DB 377 KAAEVUNAWQRNPSREVEQLLRKQSRITATNYDIYVSSAALDDILAKFSPHSALSIVIGV 436
QY 641 IVIFLYISIALGYSYSWSRVYDVKATLGLGGVAVVVLGAVMAAMGPFYSVLGIRSSSLVILQ 700
DB 437 AVTVLY---APCTLLRM-RDPVRGQSSVGVAGVLLMCPSTAAGLGLSALLGIVFNAASTQ 492
QY 701 VVPFLVLSVGNADNIPFVLEYQRLPRPCGEPREHVIHGRALGRVAPSMKLCSEATCFFL 760
DB 493 VVPFLALGLGVDHIIFMLTAAYAESNRR--EOTKL-----ILKKVGPSILFSACSTAGSFFA 546
QY 761 GALTMPAVRTFALTSLGAVILDFLLQMSAFVALLSLDSKROEASLDV-CCCV-----K 814
DB 547 AAFIPVPALKVFCLOAAIVMCSNLAAALLVFPAMISLDLRRRTAGRADIPCCCFPMWKEQ 606
QY 815 PQELPP-----PQGGEGL-----LLGFF 832
DB 607 PKVAPPVPLNNGNGARHPKSCNNRVVALPAQNPLLEQRADIPSSHSLSFSLATFA 666
QY 833 QKAYAPFLHLWITRGVVLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNR 892
DB 667 FQHYTPFLMRSWVKFLTMGFLAALISLYASTRLQDGLDIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVVTTLGYNFSSSEAGMAICSSAGCNPFSTQKIQYATFEFEQSVIATP--- 949
DB 727 LF--GFYSMYAVTQG-----NFEYPTQOQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWVDDFDIMLTPSSCCRLYISGPNKDKFCPSVTNSLNCNKNC----- 993
DB 765 KNDNGGLDPFWLLLPSEWL-----GNLQKIPDEEYRDGLTKECFPNASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PMFLNDRPNIKCPKGG 1028
DB 813 ILAYKLIQVQGHVDPNDVKELVLTNRLVNSDGIINQAFYNYLSAW-----ATNASPTL 868
QY 1029 LAA--YSTSVNLTSDGOVLASRFMAYHKPLKNSQD-----YTEALRAAR 1070
DB 869 LRANCIRNRANGASQGLYPEPRQYFHQP--NEYDLKIPKSLPLVYAQMPFYLHGLTDT 926
QY 1071 ELAAMITADLRKVPGTDPAFEV--PPYTTITNVFYQYLTILPEGLFMLSCLVPTPAVSC 1128
DB 927 QI-KTLIGHIRDLVKYEGFGLPNTPSGIPFPFWEQYMT-LRSSLAMILACVLAALVLV 984
QY 1129 LLLGLDLRSGLNLLSIVMILVDVTGFMALMDISYNAVSLINLNAVSGMSVBFVSHITRS 1188
DB 985 SLLLSVWAAVLVILSVLASLAQIFGAMTLLIGIKLSAIPAVILILSVGMWLCF--NVLIS 1042
QY 1189 FAISTKPTWLERAKEATISMGSAVAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLLITL 1248
DB 1043 LGFMTSVGNRQRVOLSMQMSLGLPLVHGLMTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102


```

; TELFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-918-658-6

Query Match          6.3%; Score 438.5; DB 2; Length 1285;
Best Local Similarity 20.0%; Pred No. 1.6e-31;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KGTSLSDKLSFTSTHTLLGQFGMTWVASWPLTILVLSVTPVVALAAGLVFTLTTDPV 380
Db 321 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 43 KKGARGRTAIVLRVFSQSHLETGLSSVQKAGKVLFAILLVSTFCVGLKSAQIHSKVH 102
Db 43 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 381 ELWAPNSQARSEKAFHQHFG-PFFRTNOVILTA---PNBSSRYDLSLLGPKNFGIL 436
Db 381 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 103 QLWIOEGGRLEAEATYKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
Db 103 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 437 DLDLLELELELERLHLQVSPQAQRNISLQDICYAPLNPNTSLY-----DCC 486
Db 437 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 154 HLEVLKATAVKVHLYDTE-W-----GLRDMCMFSTPSFEGIYIEQILRHLPFS 204
Db 154 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 487 INSLAQYFQN-----NRTLLLTAN-----QTLMGQTSQVDMKH 521
Db 487 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 IITPLDCFWEGSQLGPESAVVPGINQLRLTTLNPNASVMQMKQKMKSEKISDFETV 264
Db 205 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 522 FLY-----C-----ANAPLTFKD-----GTALALSCMADYGPVPFP 553
Db 522 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 EQMKRAAIGSGYMEKPCPLNPNCPDTPAPKNSTQPPDVGAILLSGGCYG-YAAKHMH 323
Db 265 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 554 ---LAIGYK---GKDYSAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
Db 554 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 PRELIVGGRKRRNSHLRAQALQSVQLMTEKEMVDQWQDNKYVHHLGWQEKAEVLN 383
Db 324 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 591 LWEEAF---LEEMRAFORMAGMQVFTFAERSLDEINRTTAEDLPATSYIVIFLYI 647
Db 591 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 AMQNPFSREVEQLLRKQRIATNYDIYVPSAALDILAKFSPHSALSIVIGAVTVLY- 442
Db 384 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 648 SLALGSYSSWSRVMDSKATLGLGVAVLGAVMAAGFFSYLGI----- 692
Db 648 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 443 ---AFCTLLRW-RDPVRGQSSVGAVGLLMCFSTAAGLGLUSALLGVFNALTAAYAESNR 499
Db 443 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 693 -RSSLVI---LQVVPFLVLSGADNIFTFVLEYQRLPRRPGEPREVHIGRALGRVAPSM 747
Db 693 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 500 EQTKLILKNASTQVVPFLALGLGVDFI-----VGPSI 533
Db 500 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 748 LLCSLSEAICTFFGALTTPMNAVRTFALTSGLAVILDFLQMAFVALLSDSKRQESRL 807
Db 748 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 534 LFSACSTAGSFFAAATIPVPAKVFCLQAAIVMCSNLAAALLVFPAMISLDRRTAGRA 593
Db 534 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 808 DV-CCCV-----KPEQLPP-----PGQ 823
Db 808 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 594 DIFCCCFPVWKEQKVPVPLPLNNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGS 653
Db 594 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 824 GEGE-----LILGFPQKAYAPFLHWTIRGVVLLILFLALFGVSLYSMCHISVGLDQELALPK 879
Db 824 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 654 SHLSAFSLATFAFQHYTFPLMRSWKFLTVMGFLAALISSLYASTRLQGLDIIIDLVK 713
Db 654 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 880 DSYLLDYFLFLNRYFEVGAPVYFTTLGVNFSSEAGNNAICSSAGCNFSFTQKIYATE 939
Db 880 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 714 DSNEHKFLDAQTRLF--GFYSMTAVTQG-----NFEYPTQQQLLRD 752
Db 714 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 940 FPEQSYLAIP-----ASSWVDDFDIMLTSPSSCCRLIYISGPNKDFCPSVNSLNC 989
Db 940 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 753 YHDS--FRVPHVTKNNNGGLPDPFWLLFSEWL-----GNLQKLFDEYRDGL 798
Db 753 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 990 LKNC-----MSITMGSVRPSVEQ-----PHYL-PWF 1015
Db 990 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 799 TKECWFPNASSDAILAYKLIQVOTGHVDNPVDKELVLTNRLVNSDGIINQRAFNYLSAWA 858
QY 1016 LNDRPNIKCPKGGLAAYSTSVNLTSGOVLASRFMAYHKPKNSQD----- 1061
Db 859 TND-----VFAYG-----ASQGLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQM 901
QY 1062 --YTEALARARELAANITADLRKVPGTDPFAFEV--PPYTTNNVFEQYLTILPEGLFMLS 1117
Db 902 PFYHLGLTDTSQI-KTLIGHIRDLSVKYEGFGLPNYPGIPFIWEQYMT-LRSSLAMIL 959
QY 1118 LCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFPMALWDISYNVSLINLVSAVGM 1177
Db 960 ACVLLAALVLVSLLLSVMAAVILSVLASLAQIFGAMTLLGLKLSAIPAVILILVGM 1019
QY 1178 SVRFVSHITSFATSTKPTWLERAKEATISGSAVFAGVAMTNLPGLILVLGLAKAQIQL 1237
Db 1020 MLCF--NVLLSLGFMSTVGNRRQVQLSMQMSLGPLVHGLMTSGVAVFMLSTSPFVIR 1077
QY 1238 FFRFLNLITLLGLLHGLVFLPVLVSYPGDPVNPALALEOKRAEEAAVAAVWVASCNHP 1297
Db 1078 HFCWLLLVLCVGACNSLLVPFILLSWVGPEAE-LVPLE-----HPD 1118
QY 1298 RVSTADNIYNHS 1310
Db 1119 RISTPSPLPVRS 1131

RESULT 57
US-09-724-631-6
; Sequence 6, Application US/09724631
; Patent No. 6551782
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/724,631
; FILING DATE: 28-No. 6551782-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE: 1996-05-31
; APPLICATION NUMBER: 08/540,406
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

US-09-724-631-6

Query Match 6.3%; Score 438.5; DB 2; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.6e-31;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVSVIPVVALAAGLVFTELTDPV 380
DB 43 KGKAGSRRTAIIYLRVSVFQSHLETGSSVQKAGKGVLFVAILVLSTFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSSEKAFHDOHFG-PPFRTNOVILTA---PNRSSRYVDSLLGPKNPSGIL 436
DB 103 QLMIOGGRLEBLATYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQRLHQLVMSPEAQNISLQDICVAPLNPNNTSLY-----DCC 486
DB 154 HLEVLKATAVKVLHYDTB-W-----GLRDMCNMPTSPFEGIIYIEQILRHLPSCS 204
QY 487 INSLLOYFON-----NRYLLLLTAN-----OTLMGQTSQVDWKKH 521
DB 205 IITPLDCFWGSQLGPGSAVPIPLNORLLMTLNPASVMQYMKQKMSSEKISDFETV 264
QY 522 FLY-----C-----ANAPLTKD-----GTALALSCWADYGAVPFP 553
DB 265 EQYMGAAIGSGVMKPCPLNPNPCDPDPAPKNSTQPPDVGAILLGGCYG-YAAKHMW 323
QY 554 ---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
DB 324 PEELIVGGRKRNRSGLHRAQALQSVVQLMTEKEMVDQMDNVKVHHLGWTQEKAAEVLN 383
QY 591 LWEAF---LEEMRPFQRMAGMVFQVTFABRSLEDEINRTTAEDLPFATSVYVFLYI 647
DB 384 AWORNEFSREVEQLLRQSRATNYDIYVFSAAALDDILAKFSPSALSIVIGVAVTVLY- 442
QY 648 SLALGSYSWSRWMDSKATLGLGGVAVLGVAVLGAAMWGFPSYLI----- 692
DB 443 ---AFCTLLRW-RDPVRGQSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYESNR 499
QY 693 -RSSLVI---LQVPEFLVLSGADNIFIFVLEYQRLPRRPGEPREHVHIGRALGRVAPSM 747
DB 500 EOTKLILKNASTQVFPFALGLGVDFI-----VGPFI 533
QY 748 LICSISEATCFIFGALTPMPAVRTFALTSGLAVILDFLQMGAFVALLSLDSKROBASRL 807
DB 534 LFSACSTAGSFFAAAFIPVPAKVFCLQAAIWMCSNLAALLVFPAMISLDRRRRTAGRA 593
QY 808 DV-CCCV-----XPQELPP-----PCQ 823
DB 594 DIFCCCFVWKEQPKVAPPVPLPLNNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGS 653
QY 824 GEGL-----LLGPFQKAYAPFLHWTIRGVVLLFLFALFGVLSYSMCHISVGLDQELALPK 879
DB 654 SHLSASFSLATFAFQHYTFELMSWKFTVMGFLAALISSLYASTRLQDGLDIDLVPK 713
QY 880 DSYLLDYFLNRYFVPGAPVYVFTTLGYNFSSEAGMNAICSSAGCANNFSFTQKIYATE 939
DB 714 DSNHKKFLDAQTRLF-GFSMVAVTQG-----NFEYPTQQQLLRD 752
QY 940 PFEQSLAIP-----ASSWDDPDIDMLTPSSCCRLYISGPNKDFCEPSTVNSLNC 989
DB 753 YHDS---FRVPHVIKNDGGLPFWLLLFSEWL-----GNLQKIFDEYRDGRL 798
QY 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
DB 799 TKECFPNASSDAILAYKLIIVQTHVDNPDVDELVLNRLVNSDGLIINORAFYNLSAWA 858
QY 1016 LNDREPNIKCPKGLAAYSTSVNLTSQGVLASRFMAVYHKLKNSQD----- 1061
DB 859 TND-----VFAYG-----ASQGLYPEPRQYFHQP---NEYDLKIPKSLPLVYAQM 901
QY 1062 ---YTEALRAARELANITADLRKVPCTDPAFV---PPYITNVFVEQYVITLPEGLFMLS 1117
DB 902 PFFYHGLTDSQI-KTLIGHIRDLSVKYEGFGLPNYPSPGIPFIWEQYMT-LRSSLAMIL 959

RESULT 58

US-08-954-701A-6
; Sequence 6, Application US/08954701A
; Patent No. 6610507
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-701A-6

Query Match 6.3%; Score 438.5; DB 2; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.6e-31;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVSVIPVVALAAGLVFTELTDPV 380
DB 43 KGKAGSRRTAIIYLRVSVFQSHLETGSSVQKAGKGVLFVAILVLSTFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSSEKAFHDOHFG-PPFRTNOVILTA---PNRSSRYVDSLLGPKNPSGIL 436
DB 103 QLMIOGGRLEBLATYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQRLHQLVMSPEAQNISLQDICVAPLNPNNTSLY-----DCC 486

Db	154	HLEVLVKATAVKVHLYDTE-W-----GLRDMCMPESTPSEGIYVIEQILRHLP	204
QY	487	INSLQYFON-----NRTLILLTAN-----QTLMGQTSQVDWKDH	521
Db	205	IITPLDCFWEGSQLGPESAVVIPGLNQRLMTTLNPASVMQYMKQKMBEIKISFDFTV	264
QY	522	FLY-----C-----ANAPLTFKD-----GTALALSCMADYGA	553
Db	265	EQYMKRAAIGSGYMEKPCPLNPLNCPDPAFNKNSQPPDVGAILSGGCYG-YAAKHMHW	323
QY	554	---LAIGVK---GKDYGEAEALIMTFSL-----NNYPAGDPRLAQAK-----	590
Db	324	PEELVGRKRNRSRHLKRAQALQSVQLMTEKEMYDQNDQNYKVHHLGWTQEKAAEVLN	383
QY	591	LWEEAF---LEEMRAFQRMAGMFOVTTAERSLEDEINRTAEDLPFATSYIVIFYI	647
Db	384	AWORNFREVEQLLRKQSRATNYDIYVFSAAALDDILAKFSHPALSIVIGVAVTVLY-	442
QY	648	SLALGSYSNSRWVDSKATLGLGVAVVLGAVMAAMGFFSYLGI-----	692
Db	443	--AFCTLLRW-RDPVRGSGSSVGAVGLLMCFSTAAGLGLSALLGIVFNALTAAVAESNRR	499
QY	693	---RSSIVI---LQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRALGRVAPSM	747
Db	500	EQTLLKNAQVPPFLALGLGVDFHI-----VGFPSI	533
QY	748	LCLSLSEATCFPLGALTMPAVRTFALTSGLAVIDLFLQMGAFVALLSLSKQOEASRL	807
Db	534	LFSACSTAGSFFAAAFIPVPALKVFCLOAAIWMCSNLAAALVFPAMISLDLRRRTAGRA	593
QY	808	DV-CCCV-----KPQELPP-----PQO	823
Db	594	DIFCCCFPVWKQPKVAPVPLPNNNGRGARHPKSCNNRVLPAPQPLLEQRADIFQS	653
QY	824	GEGL-----LLGFQKAYAPFLHMTIRGVVLLFLALFGVSLYSMCHISVGLDQBLALPK	879
Db	654	SHSLASFSLATFAFHYTFFLRSWVKFTVMGFLAALISSIYASTRLQDGLDIIDLVPK	713
QY	880	DSYLLDYFLFNRYFVGVAPVFTVTLGVNFSEAGMNAICSSAGCNFPSTQKIQYATE	939
Db	714	DSNEHKFLDAQTRLF--GFYSYVAVTQG-----NFEYPTQOQLLRD	752
QY	940	FPEQSVLAP-----ASWVDDFIDMLTPSCCRLYISGNKDKFCFSTVNSLNC	989
Db	753	YHDS--FRVPHVIKNDNGGLPDFWLLLFSEWL-----GNLQKIFDEEYRDGRL	798
QY	990	LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF	1015
Db	799	TKECWFPNASSDAILAYKLIQVGHVDNPVDKELVLTNRLVNSDGIINQRAFNVYLSAWA	858
QY	1016	LNDRPNIKCPKGLAAYSTSVNITSDQGVASRFMAYHKLKNSQD-----	1061
Db	859	TND-----VFAYG-----ASQGLYPEPROQYFHP--NEYDKIPKSLPLVYAQM	901
QY	1062	--YTEALRAARELANITADLRKVPCTDPAFEV--PPYITVNVYEQVLTILPEGLFMLS	1117
Db	902	PFLHGLTDSQI-KTLIGHIRDLSVKYEGFGLNTPSGIPPIFWEQYMT-LASSLAMIL	959
QY	1118	LCLVPTFAVSCLLGLDLASGLNLLSIYVILVDVTVGFMALNDISYNAVSLINLSAVGM	1177
Db	960	ACVLLAALVLVLSLLLSVAAVLVLSVLAIAQIFGAMTLLGILKLSAIPAVILLISVGM	1019
QY	1178	SVFVSHITRSPATSKPTWLERAKATISMGSAGFAGVAMTNLPGILVLGLAKAQI	1237
Db	1020	MLCF--NVLLISLGFMTSVGNRRQVRQLSMQMSGLPLVHGMTLTSGVAVFMLSTPPEFVIR	1077
QY	1238	FFERMLLITLLGLHGLFPLVLSVGVDPNPALALEQKRAEEAAVAAVMVASCNHP	1297
Db	1078	HFCWLLLVLCVGCACNSLIVFPILLMSVGEAE-LVPLE-----HPD	1118
QY	1298	RVSTADNIYVNH	1310

Db

1119

RISTPSPLPVRSS

1131

RESULT 59

US-09-754-032-6

; Sequence 6, Application US/09754032

; Patent No. 6921646

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GOORICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fiehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/754,032

; FILING DATE: 03-Jan-2001

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/540,406

; FILING DATE: 06-OCT-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1285 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-754-032-6

Query Match

Best Local Similarity

Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY

321

KGTSLSDKLSSTHTLLQFFQCGWGTWASWPLTILVLSVIPVVALAAGLVTELTDPV

380

Db

43

KGKARGRTAIYLRVSFQSHLETLSGSVQKAGKVLFAILLVLTFCVGLKSAQIH

102

QY

381

ELWAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSYRYDSLLLGPKNPSGIL

436

Db

103

QWIIQEGGRLEALAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A

153

QY

437

DLDLLELLELLEQLERHLRQVMSPEAQRNISLQDICVAPLNPNSTLSY-----DCC

486

Db

154

HLEVLVKATAVKVHLYDTE-W-----GLRDMCMPESTPSEGIYVIEQILRHLP

204

QY

487

INSLQYFON-----NRTLILLTAN-----QTLMGQTSQVDWKDH

521

Db

205

IITPLDCFWEGSQLGPESAVVIPGLNQRLMTTLNPASVMQYMKQKMBEIKISFDFTV

264

QY

522

FLY-----C-----ANAPLTFKD-----GTALALSCMADYGA

553

Db

265

EQYMKRAAIGSGYMEKPCPLNPLNCPDPAFNKNSQPPDVGAILSGGCYG-YAAKHMHW

323

QY 554 ---LAIGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----590
Db 324 PEELIVGGRKRNRSGHLRKAQALQSVQLMTEKEMDQMDQYKVHHLGWTQEKAAEVLN 383
QY 591 LWEAF---LEEMRAFQRMAGMFOVFTFAERLSLEDEINRTTAEDLPFATSYVIFLYI 647
Db 384 AQORNSREVEQLLRKQRIATNYDIYVSSAALDILAKFSPSALSIVIGVAVTVLY- 442
QY 648 SLALGYSYNSRVMYDSKATLGLGVAVVLGVMAAMGFFSYLGI-----692
Db 443 --AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI-----LQVVPFLVLSVGADNIFIVLEYQRLPRPGEPRVHIGRALGVAPSM 747
Db 500 EOTKLILKNASTQVVPFLALGLVDHIFI-----VGP81 533
QY 748 LLCSEALCFPLGALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKROEASRL 807
Db 534 LFSACSTAGSFFAAAFIPVPALKVFCLOAAIWMCSNLAALVFPAMISLDLRRRTAGRA 593
QY 808 DV-CCCV-----KPQELPP-----PGQ 823
Db 594 DIFCCCFPVWKEQKVPVPLPLNNNGRGARHPKSCNNRVPLPAQNPLLBORADIPGS 653
QY 824 GEGL-----LLGFQKAYAPFLHMTTRGVLLITRGLVGLVLSYMSCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFOHYTFPLMRSMVKFLTWGFLAALISSIYASTRLQDGLDIDLVPK 713
QY 880 DSYLLDYFLFLNRYPEVGAPVYVFTTLGYNFSSEAGMNAICSSAGCNPFSFTQKIYATE 939
Db 714 DNEHKFLDAQTELP--GFSYMAVTOG-----NFEYPTQOQLLAD 752
QY 940 FPEQSYLAIP-----ASSWDDFDIDWLTTPSSCCRLYISGPNKOKFCPSTVNSLMC 989
Db 753 YHDS--FRVPHVINDNGGLPDFWLLLFSEWL-----GNLQKIFDEEYDGR 798
QY 990 LKNC-----MSITGVSVPSEVQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLIVQTHVDNFDVKELVLTNRNVNSDGIINORAFNYLSAWA 858
QY 1016 LNDPRNICKPKGLAAYSVNLTSDOVLASRFMAYHKPLKNSOD-----1061
Db 859 TND-----VFAYG-----AQGLYPERQYFHQP--NEYDLKIPKSLPLVYAQM 901
QY 1062 --YTEALRAARELANITADLRKVPCTDPAFV--FPYITTVFVEQYLTILPEGLFMLS 1117
Db 902 PFLHGLTDSQI-KTLIGHIRDLSVYEGFGLPNVPSGIPFIFWEQYWT-LRSSLAMIL 959
QY 1118 LCLVPTFAVSCLLGLDRLSGLNLLISVIMILVDTVGFMAWDISYNAVSLINLSAVGM 1177
Db 960 ACVLLAALVLVSLLLSVMAAVLVLVSLASLAQIFGAMTLLIGIKLSAIPAVILISVGM 1019
QY 1178 SVEFVSHITRSPAISTKPTWLERAKEATISMSGSAVPAGVAMTNLPGILVLGLAKAQLIQI 1237
Db 1020 MLCF--NVLISGFTSVGNRRVQLSGMQLSLGVLPHGMLTSGVAVFMLSSTPPEFVIR 1077
QY 1238 FFFRLMLLITLGLLHGLVFLPVLVILSYVGPDPVNPALALEOKRAEBAVAVMVASCPNHP 1297
Db 1078 HFCWLLLVLCVGCNLSLVFFILLSMVGPEAE-LVPLE-----HPD 1118
QY 1298 RVSTADNIYVNSH 1310
Db 1119 RISTPSPLVRSS 1131

RESULT 60

US-08-916-140-6

; Sequence 6, Application US/08916140

; Patent No. 6946257

; GENERAL INFORMATION:

; APPLICANT: Scott Mathew P.

; APPLICANT: Goodrich, Lisa V.

; APPLICANT: Johnson, Ronald L.

; APPLICANT: Epstein, Ervin Jr.
; TITLE OF INVENTION: PATCHED GENES AND USES RELATED THERETO
; FILE REFERENCE: CIBT-P04-203
; CURRENT APPLICATION NUMBER: US/08/916,140
; CURRENT FILING DATE: 1997-08-21
; PRIOR APPLICATION NUMBER: US 08/656,055
; PRIOR FILING DATE: 1996-05-31
; PRIOR APPLICATION NUMBER: US 08/540,406
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: US 08/319,745
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1285
; TYPE: PRT
; ORGANISM: Drosophila
; US-08-916-140-6

Query Match 6.3%; Score 438.5; DB 2; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.6e-31;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KTSLSDKLSPETHLLGQFFQCGWGTWASWPLTILVLSVIPVVALAAGLVFTELTPDV 380
Db 43 KKGARSGRTAIYLRVSFQSHLETLGSSVQKHAGKVLFAVLVLSFCVGLKSAQIHKVH 102
QY 381 ELWASPNQARSEKAPHQHFQ-PFRTNOVILTA---PNRSRYRDSILLGPKNFSGIL 436
Db 103 QWIOEGGRLEALAYTOKTIGEDSSATHQLLIQTTHDPNASVLHPQALL-----A 153
QY 437 LIDLLELLELQERLHLQVMSPEAQRNISLQDICYAPLNPNTSLY-----DCC 486
Db 154 HLEVLVATAVKAVHLYDTE-W-----GLRDMCNMPTSPFEGIYIIEQLRHLIPCS 204
QY 487 INSLLOYFQN-----NRTLLLLLTAN-----QTLMGQTSQVDWMDKH 521
Db 205 IITPLDCFEGSQLGPESAVVIPGLNORLLTTLNPASVMQYMKQKSEKISFDPETV 264
QY 522 FLY-----C-ANAPLTFKD-----GTALALSCWADYGAPVFPF 553
Db 265 EQYMKRAAIGSGYMEKPCPLNPNPCPDTPAPKNSTQPPDVGAIISSGCGY-YAAKMHM 323
QY 554 ---LAIGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----590
Db 324 PEELIVGGRKRNRSGHLRKAQALQSVQLMTEKEMDQMDQYKVHHLGWTQEKAAEVLN 383
QY 591 LWEAF---LEEMRAFQRMAGMFOVFTFAERLSLEDEINRTTAEDLPFATSYVIFLYI 647
Db 384 AQORNSREVEQLLRKQRIATNYDIYVSSAALDILAKFSPSALSIVIGVAVTVLY- 442
QY 648 SLALGYSYNSRVMYDSKATLGLGVAVVLGVMAAMGFFSYLGI-----692
Db 443 --AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI-----LQVVPFLVLSVGADNIFIVLEYQRLPRPGEPRVHIGRALGVAPSM 747
Db 500 EOTKLILKNASTQVVPFLALGLVDHIFI-----VGP81 533
QY 748 LLCSEALCFPLGALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKROEASRL 807
Db 534 LFSACSTAGSFFAAAFIPVPALKVFCLOAAIWMCSNLAALVFPAMISLDLRRRTAGRA 593
QY 808 DV-CCCV-----KPQELPP-----PGQ 823
Db 594 DIFCCCFPVWKEQKVPVPLPLNNNGRGARHPKSCNNRVPLPAQNPLLBORADIPGS 653
QY 824 GEGL-----LLGFQKAYAPFLHMTTRGVLLITRGLVGLVLSYMSCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFOHYTFPLMRSMVKFLTWGFLAALISSIYASTRLQDGLDIDLVPK 713
QY 880 DSYLLDYFLFLNRYPEVGAPVYVFTTLGYNFSSEAGMNAICSSAGCNPFSFTQKIYATE 939

Db 714 DSNEHKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQQQLLRD 752
Qy 940 PFEQSYLAIP-----ASSWDDFIDMLTPSSCCRLXISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNDGGLPDFWLLLFSEWL-----GNLQKIFDEEYRDGRL 798
Qy 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLI VQTGHVDPVKELVLTNRLVNSDGINQRAFNYLSAWA 858
Qy 1016 LNDRENIKCPKGLAAYSTVNLTSDDGOVLASRFMAYHKPLKNSOD----- 1061
Db 859 TND-----VFAYG-----ASQKLYPEPRQYFHQP--NEYDLKIPKSLPLVYAQM 901
Qy 1062 --YTEALRAARELANITADLRKVCTDPAFEV--FPYTITNVFEQYLTILPEGLFMLS 1117
Db 902 PFYHGLTDTSQI-KTLIGHIRDLVKYEGFGLPNYPGIPPIFEWQYMT-LRSSLAMIL 959
Qy 1118 LCLVPTFAVSCLLGLDLRSGLNLISIVMLVDTVGFMALWDISYNAVSLINLVSAGM 1177
Db 960 ACVLLAALVLSLLLSVMAAVLVILSVLASLAQIFGAWTLGIGKLSAIPAVILISVGM 1019
Qy 1178 SVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQI 1237
Db 1020 MLCF--NVLISLGFMTSGNRQRRVQLSNQMSLGLPLVHGMLTSGVAVFMLSPPPEFVIR 1077
Qy 1238 PFRNLNLTLLGLHLGVLPVILSYGPDVNPALALEQKRAEAAVMAVASCNPNS 1297
Db 1078 HFCWLLVLCVGCAGNSLLVFPILLSMVGPEAE-LVPLE-----HPD 1118
Qy 1298 RUSTADNIVNHS 1310
Db 1119 RISTPSLPVRSS 1131

RESULT 61

PCT-US95-13233-6

; Sequence 6, Application PC/TUS9513233

; GENERAL INFORMATION:

; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/13233

; FILING DATE: 06-OCT-1990

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1285 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US95-13233-6

Query Match 6.3%; Score 438.5; DB 4; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.6e-31;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;
Qy 321 KGTSLSKLSPSTHTLLGQFFQMGTVWASWPLTILVLSPVVALAAGLVFTELTTDPV 380
Db 43 KKGARGSRTAIIYLSVFQSHLETGLSSVQKHAGKVLFAILLVSTFCVGLKSAQIHSKVH 102
Qy 381 ELWAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSYRYSLLILGPKNFSIL 436
Db 103 QLMWQEGRLAEALAYTKTIGEDESATQQLIQTTHDPNASVLHPQALL-----A 153
Qy 437 DLDLLLELLEQLERLHLQVWSPAQRNISLQDICYAPLNPDNTSLY-----DCC 486
Db 154 HLEVLVKTATVKKVHLYDE-W-----GLRDCNMPSTFEGIYIEQILRLHLFCS 204
Qy 487 INSLLOYFQN-----NRTLILLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFWEGSQLGPESAVVIPGLNQLRLWTLTNPASVMQYMKQKMSSEKISPDFETV 264
Qy 522 FLY-----C-----ANAPLTFKD-----GTALALSCMADYAGVPFPF 553
Db 265 EQYMKRAAIGSYMEKFCNLPNCPDPAFNKNSQTPPDVGAILSGCYG-YAAKHMHW 323
Qy 554 ---LAIGCYK--GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
Db 324 PEELIVGGRKNRSGHLKIAQLQSVVOLMTEKEMYDQWQNYKVHHLGWTQEKAAEVLN 383
Qy 591 LWEEAF---LEEMRAFORRMAGMFQVTFRTAERSLEDEINRTTAEIDLPIFATSYIVIFYI 647
Db 384 AWQNFSEVQELLRKQSRATNYDIVVFSSAALDDILAKFSPHSALSIVIGVAVTVLY- 442
Qy 648 SLALGSYSSWSRVWVDSKATILGLGVAVVLGAVMAAMGFFSYLGI----- 692
Db 443 --AFCTLLRW-RDPVRGQSSVGAVGLMCFSTAAGLGLSALLGIVFNALTAAAYAESNRR 499
Qy 693 -RSSIVI---LQVVPRLVLSVGADNIFI FVLEYQRLPRRPGEPREVHIGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVPFALGLGVDDHIFI-----VGRSI 533
Qy 748 LLCSLSEACFPFLGALTMPAVRTFALTSGLAVIDLFLQMSAFVALLSLSKRQKQASRL 807
Db 534 LFSACSTAGSFFAAAFIPVPALKVFCLOAAIVMCSNLAAALIVFPAMISLDLRRRTAGRA 593
Qy 808 DV-CCCV-----KPOELPP-----PQ 823
Db 594 DIFCCCFPVWKEQPKVAPFVPLNANNNGRGARHPKSCNNRVLPFPAQNPILLEQRADIPGS 653
Qy 824 GEGL---LLGFFOKAYAPFLLHWITRGVLLFLFALFGVSLYSMCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFQHYTFPLMRSWVKFLTWGFLAALISSLYASTRLQDGLDIDLVPK 713
Qy 880 DSYLLDYFLNLNRYPEVGAPVYVTTGLNPFSSSEAGMNAICSSACGNPFSTQKIQYATE 939
Db 714 DSNEHKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQQQLLRD 752
Qy 940 PFEQSYLAIP-----ASSWDDFIDMLTPSSCCRLXISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNDGGLPDFWLLLFSEWL-----GNLQKIFDEEYRDGRL 798
Qy 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLI VQTGHVDPVKELVLTNRLVNSDGINQRAFNYLSAWA 858
Qy 1016 LNDRENIKCPKGLAAYSTVNLTSDDGOVLASRFMAYHKPLKNSOD----- 1061
Db 859 TND-----VFAYG-----ASQKLYPEPRQYFHQP--NEYDLKIPKSLPLVYAQM 901
Qy 1062 --YTEALRAARELANITADLRKVCTDPAFEV--FPYTITNVFEQYLTILPEGLFMLS 1117
Db 902 PFYHGLTDTSQI-KTLIGHIRDLVKYEGFGLPNYPGIPPIFEWQYMT-LRSSLAMIL 959

US-09-332-522E-6

Query Match 2.2%; Score 150; DB 2; Length 1237;
Best Local Similarity 19.3%; Pred. No. 0.00037;
Matches 120; Conservative 86; Mismatches 240; Indels 176; Gaps 21;
QY 340 FQGGGTWVA-----SWPL-----TILVLSVIPWALAAGLVFTELT-----377
DB 10 FSKAWKFWPACGKGTFFGYFLINLPCTTPTKLVVPEYFGSLSLWHSNTSTTPOE 69
QY 378 -----DPVELWSAPNSQARSEKAFHQHFGPFFRTNQVIL-----TAPNRSSRYDS 424
DB 70 PHPGEPWPPEQVLSNSTDSPPLLPWAQSSPAFFVQIILRTSVLPWTEGMQLMD 129
QY 425 LLLGPNFSGILDLDLLELELQERLRLHLQWSPQAORNISLQDICVAPLNPONTSLYD 484
DB 130 AFRAP-----LHEVFKLLEIVRNHQ--SSENKR--TLEHNC--LHVDNVK--R 169
QY 485 CCINSLLQYFQNNRTLLLLLTANOTLMGQTSQVDWKDHLFCANAPLTFKDGTAALASCM 544
DB 170 GTHGQLDQIFPEYGCILLSPAN-----LWQNSQNFTRDNLN-----208
QY 545 DYGAFFPFLAIGGYKGDYSAEALIMTFSLNYPAG---DPLAQAKLMEEF-----596
DB 209 -----TIFQVHNLQKSKVSAEMLFGLPMQDTGFKRYPLKARSRIIQYALTFLK 258
QY 597 ---LEEMRAFQRMAGMFQVTFATRSLEDEINRT-----TAEDLPFATSIVIVFL 645
DB 259 HNDMEYLDLTKELKLRHPPLPLASASABEPTTITFYPGEYRMMELVPYTVAPMLVFA 318
QY 646 YISALGSSVSSRWVW--DSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPF 704
DB 319 YV-----YFSVRKIDVRSRFLALCSVITTAGSLANSLGCLFPFGTISLQSKDIFPY 372
QY 705 LVLSVGADNIFIVLEYQLRPPRGPPEVHIGRALGRVAPSMMLCSISEAICFFLGALT 764
DB 373 LVILVLENSLVTIKSVVSMDET--FDVKIRVAQALSKEGWHISKTLTTEITILTIGLAT 430
QY 765 PMPAVRTFALTSGLAVIDFLQMSAFVALLSLSKQOE-----GLLLG-----830
DB 431 FVPVQIEFCIFAIVGLSDFMLQMLLFSTILAMNKRTEYTABAKHLPMLLSCCTQAGR 490
QY 807 LDVCCVKPQELPP--PGQGE-----830
DB 491 QDRFREGAAPALPPFVGTFFQSRQSHPKLCPADPASVSDTSVLNVGHSPEQRIKRIKI 550
QY 831 -----PFQKAYAPFLHMI 844
DB 551 VNFWARTREFFQAFIMMIVMI 572

RESULT 65

US-09-252-991A-17657
; Sequence 17657, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17657
; LENGTH: 1467
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17657

Query Match 2.1%; Score 145; DB 2; Length 1467;
Best Local Similarity 19.8%; Pred. No. 0.0015;
Matches 269; Conservative 157; Mismatches 444; Indels 488; Gaps 71;
QY 176 RVPAAATLAVGTWCGVYGSALCNQRLNFGDTGNGLAPLDITFHLLEPQOAVSGI--233
DB 303 RQGRATDAGAVRQOAGDR--QASR-----AGAGRRP-----FHRHRPQOALCAGRRR 349
QY 234 -----QPLNEGVARCNESQG-----DDVATCS-----CQD 258
DB 350 RQDPVPAGTRPDGRRPAGRPAGRAHRRGPGPAGHADHTAPCGNRWHARRPVED 409
QY 259 CAASCPAIARQALDSTFYLGQMGSLVLIILCSFAVV--TILLVGRVAPADKSKM 316
DB 410 RGGRTMTFPR-----FFIDR-----PIFAVLVSLVMIGGIVS-----442
QY 317 VDPKKGTSLSKLSFSTHTLLGQFPQ--GWGTWVASWPLTILVLSVIP-----VVALAAGL 370
DB 443 -----FFQLPSEYPAVTPPTVQVTTAYPGANPDVIAQTAT 479
QY 371 VFTELTTPVELMSAPNSQARSEKAF-----HDQHFQPF-----RTNOVILTAPNR- 417
DB 480 PLEQAITG--VEGMLYMSQSATDGRMILTITFDQIHDPMAQIQVQNRVSRVLSLPEV 538
QY 418 -----SSRYDSLGLLGPKNFSGIILDLLELLELQERL-----451
DB 539 QROGVVTKTSPDILMVVHLLSPQRYDPLVI--SNVAYLQVRDELLRLPGISDVVWGA 596
QY 452 --RHLOW--SPE--AQRNLSQDICYAPLN-----PONTSLYDCCINSL--L 491
DB 597 GEYSRMLWLDPLIAARGLTAGEVIAAVREQNVQVAGAVGQAPDSTAFTVNTLGR 656
QY 492 QYFQNNRTLLLTANOTLMGQTSQVDWKDHLFCANAPLTFKDGTAALASCMADYGA 551
DB 657 SDEQFGDIIVRTGAD---GQVTRL-----RDVARIEMGGDAYALRSLLD--GEPAV 703
QY 552 PFLAI---GGYKGDYGEAEALIMTFSLNYPAGDPLAQAKLMEEFLEEMRAFQRM 608
DB 704 ALHIIQSPGAY--ALDTAESVRATVARLEGNFPAG-----LSARIA 742
QY 609 GMFQVTFATRSLEDEINRTTAEDLPFATSIVIVFLYISALGSSVSSRWVSDSKAT- 667
DB 743 --YDPTFVRASLQTV--ATTLLERAILLVVVVWLFL-----RSWRASLIPIMAVP 789
QY 668 LGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVG--ADNIFIFVLEYQRLP 725
DB 790 VSLVGTFF---AVHLMGF-----SLNTLSLFG--LVLSIGIVVDDAIVVVENVERHI 836
QY 726 RRPGEPREVHIGRALGRV--APSMMLCSISEAICF---FLGALTMPAVRTFALTSGLA 781
DB 837 EN-GEPPLOQAARRAMDEVTPIMAITSVLAADVFIPTAFLSLGQ--EFYRQFALTAISTI 894
QY 782 LDFLLQMSAFVALLSLSKQREASRLDVCCCKPQELPPGQEGELLGFFQKAYAPELL 841
DB 895 LSALESLTSPALAGL-----LLRPR---PAG---GAVAGRFOR-----LL 929
QY 842 HWITR-----GWLLFLFALFGVSLYSMSCHISVGLDQELALPKDS 881
DB 930 QVLGRPLNAPAEYGNVKKVVRVSGLAIVYVGLGLTW-----VGfq---AVPQ---977
QY 882 YLLDYFLFLNRYFEVG-----APVYFVTTLGVNFSSEAGNNAICSSAGCN 927
DB 978 ---GFVPMQDKYLVGIAQLPNGASLERTDAVVRQMSRIGLD---EPGVESVAVFPGLSV 1031
QY 928 FSTQKIQYATPEPQSYLAIPASSWVDDFDWLTPSSCCRLYISGPNKDKFC-----980
DB 1032 NGFVNVPNAAMVF-----FMDPFES-----RTSADLGAIAIAGRLQARFASIPDGFL 1079
QY 981 ----PSTVNSLNLCKNCSITMGSVRPSVEQPHKYLPMFLNDRPN-----I 1022
DB 1080 GVFPFPPVPGLG-----TIGGFMQVE-----DRGGAGLEALARQTQVLWM 1120
QY 1023 KCPKGG--LAAYSTSVNLT-----SDQVLSARFWAY-----1052

Db 1121 KATESGQLGLMTSFDINAPQLEVVVDRTKVKSGQVRLADVPFALQVYLGSLYINDPNRF 1180
QY 1053 -----HK-----PLKNSQDYTEALRAARELAAN--ITA 1078
Db 1181 GRITYKVTQAADAPHRMQAEAGELQVRNAGAMPLPSSPVTVPSSGPRVHYNGYPSA 1240
QY 1079 DLR--KVPGTDP-----AFREVP-----YTIITNRYEQYLITLPEGLFMLSLCLVP 1122
Db 1241 DISGGALPGVSSQAVAMERLAGEVLPFGMTPEWTDLTVOQKLA--GNSALFIFPLCVL- 1298
QY 1123 TPAVSCLLGLDLRSGLN---LLSIVMLIVDTVGFMAWLDISYNAVSLINLVSAGMS- 1178
Db 1299 ---LAYLILAAQYNSWLLPLAVLILVPMCLLSAIGVWLMGGDNNVFVQIGLVVLVGLAA 1355
QY 1179 -----VEFVSHITRSFAISTKPTWLERAKEATISMSAVFAGVAMTNL---PGILVLGL 1229
Db 1356 KNAILIVEPARTLEAGA-----RALEAVEACRLRLRPLMTSLAFTAGVVPFLVM 1406
QY 1230 AXA-----QLIQIFFRNLNLLTLLGLLHGLVFLPVI 1261
Db 1407 ASGAGAEQMGVAVFAGMLGVTLFGLFITPVFVYLV 1444

RESULT 66

US-09-252-991A-16989
; Sequence 16989, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16989

LENGTH: 1051

TYPE: PR1

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-16989

Query Match 2.0%; Score 140.5; DB 2; Length 1051;
Best Local Similarity 20.5%; Pred. No. 0.0022;
Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;
QY 506 NOTLMGOT--SQVDMKHFLYCANAPLTFKGTALALSCWADYGAPVFPFLAIGGYKGD 563
Db 236 NATIIGKTLQAEQENILLKVN-----PDGSQVRLKDVADV-----LG---GQD 279
QY 564 YSEAEALIMTFLSNYP-----GDPRLAQAKLWEE--AFLEEMRAFQRMAGMFO 612
Db 280 YS-----INAQNGSPASGAIKATGANALDTAKAIQTIANLEPFPQGMKVVPYD 333
QY 613 VTFTARSLEDEINTTAEADLPFATSYIVIFLYISLALGSYSSWSRVVDSKATL--GL 670
Db 334 TTPVVSASIH-EVVKTLGE-----AILLVELVMYFLQNF-----RATLIPTI 375
QY 671 GGVAVVLG--AYMAAGPFSYLGRSSLVILQVVPFLVLSVG--ADNIFIFVLEVRQLPR 726
Db 376 AVFVLLGTFFGVLAAPF-----SINTLMFG--MVAIGLLVDDAIVVENVERVMA 426
QY 727 RPG-EPREV--HIGRALGR-VAPSMLLCSLSEACFFLGAITPMPAVRFTALTSLGLAVI 781
Db 427 BEGLSPREARKSMQIQALGVIAWLVSAVFLPAFFGSGTGV--YQFSITIVSAMA 484
QY 782 LDFLLQMSAFVALLSDSKRQEARLDVCCC---VKPQELPPPGQGBGLLLGFFQKAYAP 838
Db 485 L-----SVIVALI-----LTPALCATMLKPIEKGDHKGHGFQGFQFNRMFLS 527

QY 839 FLIHWITRGV-----LILFLALFGVSLYSMCHISVGL-----DQ-----ELALPKD 880
Db 528 -TTHGVERGVASLTKHRAQYLLIYVIVAGMIWMFTRIPTAPLPDDEQGVLPFAQVQTTPG 586
QY 881 S-----YLLDYFLFLNRYFVGAPVYFVTTLLGYNFSSE----- 913
Db 587 SSAERTQVVVDVSMREYLLEK-----ESSSVSSVFTVTGTFNAGRGQSSGMAFIMLKP 638
QY 914 -----AGMNAI-----CSSAGCNFSTQKIQYATEPPEQSVLAIPASSWDDPFD- 959
Db 639 WEERPGENSVFELAKRAQMHFFSFKQAMVFA--FAPPSVLELGNATGDFLQDQAGVG 696
QY 960 ---MLTPSSCCRLYISGPNKDKFCPSITVN-----SLNCLKNCMSI 996
Db 697 HEVLLQARNKFLMLAAQNALQVRPNHGSDEPOYKLEIDDEKASALGVSLADINSTVSI 756
QY 997 TWGSV-----RPSVEQPHKYLPMFLNDRPN-----IKCPKGGLA---AYST 1034
Db 757 AMGSSYVNDFIDRGRVVR-----YLGCRPDARMNPDDLKWKYVRNDKGMVFPFNAPAT 810
QY 1035 -----SVNLTSDGQVILASRFMAYHKPLKNSQDYTEALRAARELAANITADLAKVGTDP 1088
Db 811 GKWEYGSPLERYNGVPAEILGEPAPGLSSGD---AMAAVEE---IVKQLPKGVG--- 860
QY 1089 AFEVFPYITITNRYEQYLT-----ILPEGLFMLSLC-----LVPTFAVS 1127
Db 861 -----YSWTGLSYBERLSGSOAPALYALSLIVVFLCLAAALYESWSIPFSVMLVPLGVI 914
QY 1128 CILLGLDLRSGLNLLSIVMLIVDTVGFMAWLDISYNAVSLINL-----V 1172
Db 915 GALLATSMR-GLSNDVFFQVGLLTIGLSA-----KNAILIVEPAKELHEQKGIIVEAAI 968
QY 1173 SAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMSAVFAGVAMTNLPGILVL-GLAK 1231
Db 969 EACRMRLRPVMTSLAFTILGVVP-----LAISTG-----AGSGSQAHTGTVIGMVT 1016
QY 1232 AOLIQIFPERL 1242
Db 1017 ATVLAIFWVPL 1027

RESULT 67

US-09-328-352-4877
; Sequence 4877, Application US/09328352
; Patent No. 6562958

GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4877

LENGTH: 1005

TYPE: PR1

ORGANISM: Acinetobacter baumannii

US-09-328-352-4877

Query Match 2.0%; Score 140; DB 2; Length 1005;
Best Local Similarity 20.0%; Pred. No. 0.0023;
Matches 167; Conservative 128; Mismatches 328; Indels 212; Gaps 40;

QY 569 ALIMTFLSNYPAGDPRLAQAKLWEEAFLEEMRAFQRMAGMFOVTTFTAERSL----- 622
Db 105 ALNPTY--NNSVVSQDQTA-AILL-LEESPEGFQKMMGPINKIVSEQSKDMTISVG 159
QY 623 -DEINRTTAEDLP-----IPATSYIVIFLYISLALGSYSSWSRVVDSKATLGLGVAVV 676
Db 160 GNPVYLDKAEYDKRINILFPPIAVLIVGL---LHFAFRSKQGLILPLVTAL---LAVA 212
QY 677 LGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLEY-----QRLPRRGEPR 732

Db 213 WG--MGMMGLK-----QPMDFNSFTPIILIAAGHAVQLLKRYEDFDRLIAQMEPK 266
QY 733 EVH-----IGRALGRVAPSMMLCSLSBAICFFIGALTPMPAVRFAITSLGLAVILDFLQMS 789
Db 267 AANSEAVQSLVRVGVFVMMVLAGGIAAGFFSLLTFTIPIRSFGIFTGIGIISTIVIENT 326
QY 790 AFVALLSLSKQESRLDVCCVKPQELPPG-----QEGILLGFFQKAYAPFLHMIT 845
Db 327 FIPALRSM-----LPPSVVKVRKGL-----P-IWDWIP 355
QY 846 RGV-----VLLFLALFGVSL--YSMCHISVGLDQELALPKDSYLLDYFLFLNRY 893
Db 356 NRIGDIVLSVRPMLMTAIAAGILLAIIGTSRIIVDNDNRNFFSRDLPMQDDRFILNQ- 414
QY 894 FEVGAPVYFVTTLGYNFSEAGMNAICSSAGCNSFTQKIQYATEFPEQSYLAIPASSW 953
Db 415 -----SLGCTNSL-----YIMVDTKVRDGIENPE----- 438
QY 954 VDDFDMLTPSSCCRLYISGPNKDKFCS-----TVNSLNLCKNCMSITMGSVRPSVEQ 1007
Db 439 ILKADID-----NTEKPAFNSIPEVGKTIISIVDIYKR-MNQAMNADQPAFQ 482
QY 1008 FHKYLPWFNLDRPNIKCPKGGIAAYSTVNLTSQGVLAS-----RFMAYHKPLKNSQD 1061
Db 483 -----VPATKOVVAQYLLLYSMGSEPTDFDSYIDTTQRYAKITILLKTGSN 528
QY 1062 Y--TEALRAARELANITADLRKVPCTDPAPFVFPYTTNNVFEQYVTLPEGLFMLSIC 1119
Db 529 HRIKEILESMTYAGQLGDKAVVSGGDVDTQTIALTETW-HGKLMNLI-----QISFA 582
QY 1120 LVPTAVSCLLGLDLRSGLNLLSIVMLTDTVGFMALWDISYNVSLINLVSAVMSV 1179
Db 583 V---PFISALVF-RSISAGLIVLTPLFSILAIFGVMGMLDIPNLIPNSLISAMAVGICA 638
QY 1180 EF-----VSHITSFAISTKPT---WLERAKEATISMGSAVFAGVAMTNLPGLVL-- 1227
Db 639 DYAIYFLYRLREILREEGDIDKAIKTLSTAGKASLFVATAVAGY-----GVLSLSQ 692
QY 1228 GLAKAQLIQIFFRNLTLTLLGLHGLVFLPVILSYGPDVNPALALEKQRAEBAVAV 1287
Db 693 GFHVHQLWAMFI-----VIAMLSFVATLIMVPTMLI-----LKPRFIFSSKKKSIPVAQT 744
QY 1288 MWAS-----CPNHPSRVSTADNI--YVNHSEGS-IKGAGAISNPL--PNNGRQ 1331
Db 745 VVTSLLLGTALTSWPKTSHADEVQDIVNRSDASKFLSTASAKFILTSKNGEQ 799

RESULT 68
US-09-902-540-12490
; Sequence 12490, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12490
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12490

Query Match 2.0%; Score 138; DB 2; Length 826;
Best Local Similarity 19.4%; Pred. No. 0.0025;
Matches 134; Conservative 116; Mismatches 253; Indels 187; Gaps 33;

QY 613 VTFTAERSLEDEINRTTAEDLPFATSVIVIFL-YISLALGSYSWSRVMVDSKATLGLG 671
Db 242 VTGTFKKKIDQ--RVIVGDLASASTLAWLLLAYLAFHPSALS----- 284
QY 672 GVAIVLGAVNMAAGF-FSYLGIRSSVLVQVVPFLVLSVGADNI--FIFVLEYQRLPRRP 728
Db 285 -VGLTMAPVMAGLQWTYGVGLAYQVNL-LTGFLAAVGLGLGVEHGHILLGRYTTLRSE 342
QY 729 GEPREVIHGRALGRVAPSMMLCSLSSEALCPFLGALTMPAVRFAITSLGLAVILDFLQ 788
Db 343 GMSSEEAVDSEFRHTGSSLLAALVAALTFLUSLAWSEFRAPREFGIIAAGMLVSIYSV 402
QY 789 SAFVALLSLSKQESRLDVCCVKPQELPPGQEGLLGFFQKAYAP--FLHMITR 846
Db 403 LLLPALLGLATRFQMSPRV-----QEG-----AAGPLSLLARWLPR 439
QY 847 GVVLLFLALFGVSLYSNMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTTL 906
Db 440 -----SYRGVGIV---VGVG---LALVSQAY-----RI 462
QY 907 GYNFSESAGMNAICSSAGCNSFTQKIQYATEFPEQSYLAIPASSWVD----- 955
Db 463 SFNYDSRTLEDYKQASA-----VLDQVNDILGYSQTPVVVITDSQEWEREVRQLEARK 517
QY 956 -----DFI---DMLTPS-----SCRLYISGPNKDKFCSPTVNSLNLCKNCM 994
Db 518 VARGKESTIDFVGALDVLVPRQDEKQAILQASAKLEKLDPERLPEDTRNL-----VR 572
QY 995 SITMGSVRPSVEQFHKYLPWFNLDR-PNICKPKGGIA-AYSTSVNLTSQGVLASRPMAY 1052
Db 573 ALNMAKAPFTQE---ALPTSVRHQFESLDGSTGCVLVLYAGVSL-SDGE--GTRKFA- 625
QY 1053 HKPLKNSQ--DYTEALRAARELANITADLRKVPCTDPAPFVFPYTTNNVFEQYVTLPL 1110
Db 626 -KEVRGLQMPGQSQVSAAGEAL--ILADILDM-----VSR 657
QY 1111 EGLFMLSICLVPTFAVSCLLGLDLRSGLNLLSIVMLTDTVGFMALWDISYNVSLIN 1170
Db 658 EGPRILAAVLSVLVAMWLTIG-KLRTALICMLPTLLSVGLVGLMSLLGLQFNVLNIMV 716
QY 1171 LVSAVMSVFEVSHITRSFAISTKPTWLERAKE---ATISMGSAVFAGVAMTNLPGLVL 1227
Db 717 LPVLVGTVDAGVHLVQRLG-----ERGADFVSVAETGRAIMGGL-LTSAIGFVAL 767
QY 1228 GLAK-----AQLIQIFFRNLTLTLLG 1250
Db 768 ILAKHPGLNSIGDLANTL-GFGINMNVILLG 796

RESULT 69
US-09-489-039A-10700
; Sequence 10700, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10700
; LENGTH: 1040
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10700

Query Match 2.0%; Score 137.5; DB 2; Length 1040;
Best Local Similarity 18.1%; Pred. No. 0.0042;
Matches 162; Conservative 127; Mismatches 289; Indels 315; Gaps 41;

Db 604 NIDKAVELYSSGERKLANVFDLETMP-REYNYYAQSIISGLRKDFPNSQRNDRRIIQDF 662

QY 630 AEDLPATSYIVIFLYISLALSYSSWSRVWVDSKATLGLGGAVALGAVMAAMGFSSY 689

Db 663 SEILADLGSIGKVI---VNVASGAFSLFGIVT-----GILNFIKNPLGGMFTFLL 710

QY 690 LGIRSSILVILQV 701

Db 711 IGAIVILVILV 722

RESULT 74

US-08-720-229-14

; Sequence 14, Application US/08720229

; Patent No. 6022542

; GENERAL INFORMATION:

; APPLICANT: Rose, Timothy M.

; APPLICANT: Bosch, Marnix L.

; APPLICANT: Strand, Kurt

; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV

; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES

; NUMBER OF SEQUENCES: 100

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morrison & Foerster

; STREET: 755 Page Mill Road

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/720,229

; FILING DATE: 26-SEP-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Schiff, J. Michael

; REGISTRATION NUMBER: 40,253

; REFERENCE/DOCKET NUMBER: 29938-20002.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 813-5600

; TELEFAX: (415) 494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 808 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-720-229-14

Query Match 1.9%; Score 130.5; DB 2; Length 808;

Best Local Similarity 21.3%; Pred. No. 0.013;

Matches 92; Conservative 71; Mismatches 162; Indels 107; Gaps 18;

QY 323 TSLSKLS--FSTHLLG--QFQGGTGWASWPIITLVLSVPPVVALAAGLVFTLTDD 378

Db 345 TTLEDKIRAVNTHTPNGTAYYYQTEGGMILVWQ-----PLIATIELEAMLEATTS 395

QY 379 PYVELSAPNSQARSEKAFHQDFPFRINQVILTPAPNSSRYDLSLLGPNFSGILD 438

Db 396 PVP--SAPTSRSKRKAIRSDVAGSENNVFL---SQIYAYDKLRQSIINN----- 444

QY 439 DLLLELLEL---QERLRHQVNSPRAQRN-LSLQDICTA-PLN----PNTSLYDCCINS 489

Db 445 --VLEELATWCREQVROTVMWYEAKINPTSVMTAIYKPVSRKALGDVISVTEC---- 498

QY 490 LLOYFQNNRTLLLTANQTMQTSQVDWKDHFLYCANAPLTFKO----- 534

Db 499 -----INVDOSSVSIHKSILKTENNDCVSRPVPVTFKFVNSSQLFKGQLGARN 545

QY 535 ----GTALALSQWADYGAPVFPFLAIGGYKGKDYSEAE-----ALIMTF-- 574

Db 546 EILLSESLVENC--HQNAETFTAKNETHPKNYHVHVELPVNNISTLTDTFLALNITFIE 603

QY 575 -----SLNNYPAGDPRLAQAKLWBSAFLEEMRAFORMAGMFQVTFATERSLEDEINRTT 629

Db 604 NIDKAVELYSSGERKLANVFDLETMP-REYNYYAQSIISGLRKDFPNSQRNDRRIIQDF 662

QY 630 AEDLPATSYIVIFLYISLALSYSSWSRVWVDSKATLGLGGAVALGAVMAAMGFSSY 689

Db 663 SEILADLGSIGKVI---VNVASGAFSLFGIVT-----GILNFIKNPLGGMFTFLL 710

QY 690 LGIRSSILVILQV 701

Db 711 IGAIVILVILV 722

RESULT 75

US-09-252-991A-24533

; Sequence 24533, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 24533

; LENGTH: 1069

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24533

Query Match 1.9%; Score 130.5; DB 2; Length 1069;

Best Local Similarity 21.2%; Pred. No. 0.021;

Matches 123; Conservative 80; Mismatches 208; Indels 169; Gaps 27;

QY 815 PQELPPPGGEGLLGLFPQKAVAP--FLHWTTRGVVLLFLALFGVSLYSMCHISVGLD 872

Db 19 PGAAPEPGSGG-----AAVTPRAGISGWCIRHPATALTALTASLLGLLAFRLGV- 71

QY 873 QELALPKDSYLLDYFLNRYPEVGAPVVF---VTTLGYNFPSSBAGMNAICSSAGCN-- 926

Db 72 --APLPEADPPT---IQINALLPGSGPETMASSVATPLEVQSAIPGITEMTSSSALGTT 126

QY 927 ----NFSFTQKIQYATFPFQSYLAIPASSWVDDFDMLTPSSCCRLYISGN----- 975

Db 127 TLTQLFSLDKSIDVAA--QEVQAAINAAAG-----RLPVDMDNLPTRKV 169

QY 976 -----KDXECP-----STVNSLNCCLKMCSITMCSVSRPSVEQ 1007

Db 170 NPADSPIMILRVNSEMPLIELSDYAETILARQLSQVNGVQI-----FVVGQQRPAI-- 222

QY 1008 PHKLPWFNDRPNKCPKGGLAAY-----STSVNLT-----SDQVQLASRFM 1050

Db 223 -----RLOAQPEKLAAYQLTLADLRQSLQASVNLAKALYGEGRVST---L 266

QY 1051 AYHKPLKNSQDYTEALRAARELANITADLRK-----VPG----- 1085

Db 267 AANDQLFNASDYDDLWAVRQCAPVFLKDVARIVSAPDDYVQAWPNGVPGVALVILRQP 326

QY 1086 ----TDPA-----FEVPPYTI-TNVFEQVLTILPEGLFMLSCLVTFVAVSCLL 1130

Db 327 GANIVDTADAIOAALFRLREMLPATIEVDVLNDRRTTI-RSSLHEVELTLLLTIGLVILV 385


```
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2020
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2020

Query Match      1.8%; Score 126; DB 2; Length 801;
Best Local Similarity 19.9%; Pred. No. 0.033;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 62 LNMHPHIGMNFNLYVDGLGLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 121
QY 874 ELALP-KDSYLLDY-----FLFLNRYFEVGAPVY-----FVTILGYNFSEAGM 916
Db 122 MGVVLSDFNFIILYLFWEITSSPFLISFWREKKASYCAQKSLIITVLG-GLSMLGGI 180
QY 917 NAICSSAGCNNSFTQKIQYATEPPEQSYL-----AIPASSWVDDFDW 960
Db 181 --ILLSLATDTFSIQAMISKASDIQNSPPFILLVWILFMIGATKSAQVFFYIWLDPAMEA 238
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNLKNCMSITWGSVRPSVE 1006
Db 239 PTPVSAYLHSATMVKAGLYLIARITPFAISEGWWTIT-----LVGLITLFWASLNATQ 294
QY 1007 QPHKYLWFLNDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAHYHKLKNSQDYTEAL 1066
Db 295 HDLK-----GILAFSTVSQGLMWSMLGIGAVSYHYQGANSQLYVAGF 337
QY 1067 RAARELAANI-----TADLRKVPGTDPAPFVFPYITNVFYEQVILT 1107
Db 338 VAAIPLHNLHATFKGALFMITGGIDHSTGTRDVKKLGG--LLTIMPISFT-----LT 387
QY 1108 ILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLLSVIMLVDTVGFMAWDISYNAVS 1167
Db 388 VIT-----TLSMAGVPPP--NGFLSKEKFLSMTNTHLMLSINTLG-----ILLPIIA 435
QY 1168 LNLVSAGVMSVEFVSHITRSPFAISTKPTWL-ERAKEATISMGSVAFAGVAMTNLPGLIV 1226
Db 436 IIGSITFTVYSIKFILHI---PFGSYKPEALPKQAHESSILM---LISPIILTSL--VIV 487
QY 1227 LGLAKAQIQLIFFRNLNLIITLLG-----LLHGLVFLPVLISYVG 1266
Db 488 FGLFPFSILTQSIIEPASVAVSQTSNITAEFHLFHGIT--PAFLSTIG 532

RESULT 78
US-09-134-001C-3105
; Sequence 3105, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3105
; LENGTH: 808
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3105

Query Match      1.8%; Score 126; DB 2; Length 808;
Best Local Similarity 19.9%; Pred. No. 0.034;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 69 LNMHPHIGMNFNLYVDGLGLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 128
QY 874 ELALP-KDSYLLDY-----FLFLNRYFEVGAPVY-----FVTILGYNFSEAGM 916
Db 129 MGVVLSDFNFIILYLFWEITSSPFLISFWREKKASYCAQKSLIITVLG-GLSMLGGI 187
QY 917 NAICSSAGCNNSFTQKIQYATEPPEQSYL-----AIPASSWVDDFDW 960
Db 188 --ILLSLATDTFSIQAMISKASDIQNSPPFILLVWILFMIGATKSAQVFFYIWLDPAMEA 245
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNLKNCMSITWGSVRPSVE 1006
Db 246 PTPVSAYLHSATMVKAGLYLIARITPFAISEGWWTIT-----LVGLITLFWASLNATQ 301
QY 1007 QPHKYLWFLNDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAHYHKLKNSQDYTEAL 1066
Db 302 HDLK-----GILAFSTVSQGLMWSMLGIGAVSYHYQGANSQLYVAGF 344
QY 1067 RAARELAANI-----TADLRKVPGTDPAPFVFPYITNVFYEQVILT 1107
Db 345 VAAIPLHNLHATFKGALFMITGGIDHSTGTRDVKKLGG--LLTIMPISFT-----LT 394
QY 1108 ILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLLSVIMLVDTVGFMAWDISYNAVS 1167
Db 395 VIT-----TLSMAGVPPP--NGFLSKEKFLSMTNTHLMLSINTLG-----ILLPIIA 442
QY 1168 LNLVSAGVMSVEFVSHITRSPFAISTKPTWL-ERAKEATISMGSVAFAGVAMTNLPGLIV 1226
Db 443 IIGSITFTVYSIKFILHI---PFGSYKPEALPKQAHESSILM---LISPIILTSL--VIV 494
QY 1227 LGLAKAQIQLIFFRNLNLIITLLG-----LLHGLVFLPVLISYVG 1266
Db 495 FGLFPFSILTQSIIEPASVAVSQTSNITAEFHLFHGIT--PAFLSTIG 539

RESULT 79
US-09-252-991A-27976
; Sequence 27976, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27976
; LENGTH: 1123
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27976

Query Match      1.8%; Score 125.5; DB 2; Length 1123;
Best Local Similarity 22.1%; Pred. No. 0.067;
Matches 108; Conservative 82; Mismatches 191; Indels 107; Gaps 25;

QY 409 QVILTAPNRRSYRYDSSLGKPNFSGIILDLLELLELQERLRLHQLWMSPEAQRNISIQ 468
Db 141 QLIVLTLS-SSGRYDAMDGLBIASSNV-----LQALRRVEGVGVKVTWGAETAMRIWPD 193
```

```
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2020
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2020

Query Match      1.8%; Score 126; DB 2; Length 801;
Best Local Similarity 19.9%; Pred. No. 0.033;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 62 LNMHPHIGMNFNLYVDGLGLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 121
QY 874 ELALP-KDSYLLDY-----FLFLNRYFEVGAPVY-----FVTILGYNFSEAGM 916
Db 122 MGVVLSDFNFIILYLFWEITSSPFLISFWREKKASYCAQKSLIITVLG-GLSMLGGI 180
QY 917 NAICSSAGCNNSFTQKIQYATEPPEQSYL-----AIPASSWVDDFDW 960
Db 181 --ILLSLATDTFSIQAMISKASDIQNSPPFILLVWILFMIGATKSAQVFFYIWLDPAMEA 238
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNLKNCMSITWGSVRPSVE 1006
Db 239 PTPVSAYLHSATMVKAGLYLIARITPFAISEGWWTIT-----LVGLITLFWASLNATQ 294
QY 1007 QPHKYLWFLNDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAHYHKLKNSQDYTEAL 1066
Db 295 HDLK-----GILAFSTVSQGLMWSMLGIGAVSYHYQGANSQLYVAGF 337
QY 1067 RAARELAANI-----TADLRKVPGTDPAPFVFPYITNVFYEQVILT 1107
Db 338 VAAIPLHNLHATFKGALFMITGGIDHSTGTRDVKKLGG--LLTIMPISFT-----LT 387
QY 1108 ILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLLSVIMLVDTVGFMAWDISYNAVS 1167
Db 388 VIT-----TLSMAGVPPP--NGFLSKEKFLSMTNTHLMLSINTLG-----ILLPIIA 435
QY 1168 LNLVSAGVMSVEFVSHITRSPFAISTKPTWL-ERAKEATISMGSVAFAGVAMTNLPGLIV 1226
Db 436 IIGSITFTVYSIKFILHI---PFGSYKPEALPKQAHESSILM---LISPIILTSL--VIV 487
QY 1227 LGLAKAQIQLIFFRNLNLIITLLG-----LLHGLVFLPVLISYVG 1266
Db 488 FGLFPFSILTQSIIEPASVAVSQTSNITAEFHLFHGIT--PAFLSTIG 532

RESULT 78
US-09-134-001C-3105
; Sequence 3105, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3105
; LENGTH: 808
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3105

Query Match      1.8%; Score 126; DB 2; Length 801;
Best Local Similarity 19.9%; Pred. No. 0.033;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 62 LNMHPHIGMNFNLYVDGLGLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 121
QY 874 ELALP-KDSYLLDY-----FLFLNRYFEVGAPVY-----FVTILGYNFSEAGM 916
Db 122 MGVVLSDFNFIILYLFWEITSSPFLISFWREKKASYCAQKSLIITVLG-GLSMLGGI 180
QY 917 NAICSSAGCNNSFTQKIQYATEPPEQSYL-----AIPASSWVDDFDW 960
Db 181 --ILLSLATDTFSIQAMISKASDIQNSPPFILLVWILFMIGATKSAQVFFYIWLDPAMEA 238
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNLKNCMSITWGSVRPSVE 1006
Db 239 PTPVSAYLHSATMVKAGLYLIARITPFAISEGWWTIT-----LVGLITLFWASLNATQ 294
QY 1007 QPHKYLWFLNDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAHYHKLKNSQDYTEAL 1066
Db 295 HDLK-----GILAFSTVSQGLMWSMLGIGAVSYHYQGANSQLYVAGF 337
QY 1067 RAARELAANI-----TADLRKVPGTDPAPFVFPYITNVFYEQVILT 1107
Db 338 VAAIPLHNLHATFKGALFMITGGIDHSTGTRDVKKLGG--LLTIMPISFT-----LT 387
QY 1108 ILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLLSVIMLVDTVGFMAWDISYNAVS 1167
Db 388 VIT-----TLSMAGVPPP--NGFLSKEKFLSMTNTHLMLSINTLG-----ILLPIIA 435
QY 1168 LNLVSAGVMSVEFVSHITRSPFAISTKPTWL-ERAKEATISMGSVAFAGVAMTNLPGLIV 1226
Db 436 IIGSITFTVYSIKFILHI---PFGSYKPEALPKQAHESSILM---LISPIILTSL--VIV 487
QY 1227 LGLAKAQIQLIFFRNLNLIITLLG-----LLHGLVFLPVLISYVG 1266
Db 488 FGLFPFSILTQSIIEPASVAVSQTSNITAEFHLFHGIT--PAFLSTIG 532

RESULT 78
US-09-134-001C-3105
; Sequence 3105, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3105
; LENGTH: 808
; TYPE: PRT
```

QY 469 DICVAPLNPDNTSLYDCCINSLLQYFQNNRTLL-----LLTANQTLMGQTSQVD-----WK 519
Db 194 P---AKLTSMLSASD-LVNVRHR--NARLTGVDIGNLGVDPDSAPTATVKVDDTLVTP 247
QY 520 DHFLYCANAPTEK-DGTALALSMA--DYGAPVPPFLAIGYKGKGYSEAEALIMTSL 576
Db 248 EQF---GEIPLRADGGAIRLRDVARVERFGQSEYGFV-----SRVQMTATGLA 294
QY 577 NNYPAGDPRLAQAKLEBAPLEEMRAFORMAGMFQVTTFAERSLEDEINRTTAEIDLPIF 636
Db 295 VKMAGSNAVATAKIRATLDELSRYFPEGVS--YNIPDYTSAFVEISIRKVVSTLEAM 352
QY 637 ATSYIVIFLYISLALGSSYSRVMVDSKATILGLGGVAVILGAVMAAGFFS-YLGRSS 695
Db 353 LLVFAWMLF-----MONFRATL-----IPTLVVPVALLGTFTVMVLGLGFS 393
QY 696 LVILQVFPFLVLSVG--ADNIFIFVLEYQRLPRPG-EPREVIHIGRALGRVAP-----SM 747
Db 394 INVLMFG-MVLAIGILVDDAIIVVENVERLMAEEGLSPHDATV-KAMRQISGAIVGITV 451
QY 748 LCLCSLSEAIICFGLGALTMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDSKROEASRL 807
Db 452 VLVSFVFPMAFFSGAVGNI--YRQFAVT--LAVSIGF---SAFLAL-----SLT 493
QY 808 DVCCCKVQQLPPPGQGGELLGLGFFQKAYAP-----FLLHWITRGV 848
Db 494 PALCATLRLPIDADHHEKRGFFGFENRAFLRLTGRYRNAVAGILARPWRMLVYALVIGV 553
QY 849 VLLFLAL 856
Db 554 VALLFVRL 561

RESULT 80

US-09-252-991A-28767
; Sequence 28767, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28767
; LENGTH: 789
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28767

Query Match 1.8%; Score 123.5; DB 2; Length 789;
Best Local Similarity 21.9%; Pred. No. 0.056;
Matches 154; Conservative 98; Mismatches 240; Indels 211; Gaps 41;
QY 635 IFATSVIVIFLYISLALGSSYSRVMVDSKATILGLGGVAVILGAVMAAGFFSYLGRS 694
Db 236 VMVVGFGICLLITLVL---LYWFTKCI--RSTIAV--LITTLAVVLWQLGILLNLVGFGL 288
QY 695 SLVILQVFPFLVLSV-----GADNIFIFVLEYQRLPRR-----PGEPRE 733
Db 289 DPFYSM-LVFFLFAIGISHGVQKINGIALQSSGADN-----ALMAARLTFRQLFLPG-----339
QY 734 VHIGRALGRVAPSMILCSISEAICFPLGALTMPAVRTFAL--TSGLAVIDLDFLLQMSA 790
Db 340 -----MTAILADAVGFTLLVDIGVIRELAIGASIGAVIVFTNLTLLPV 385
QY 791 FVALLSLDSKROEASRLDVCCCKVQQLPPPGQGGELLGLGFFQKAYAPFLLHWITRGVVL 850

Db 386 AISVIGISKAVQSKDDAV-----REHP-----FWRLLSNFASPKVAP-----VSI 427
QY 851 LLFLALPGVSLYSMCHISVG-LDQ---ELALPKDSYLLDY-FLFLNRYFEVGAPVYFVTT 905
Db 428 AIALMLGGGLWYKGKHLKIGLDQCAPELR-PDSRYNLNDNDFIIRN-----472
QY 906 LGYNFSSSEAGNVAI-CSSAGCNNSFTQKIQYATEFFPEQSYLAIPASSW-----VDDF 957
Db 473 --YSTSSDVLVVMVKTSPGCGST-----HQAAMDELAWKLENTGVSQA 516
QY 958 IDMLTPSSCCRLYISGNPKDFCPSTVN-SLNCCLKNCMSITWG-----SVRPSVEQFHK 1010
Db 517 ISMTVVS---KQVIGKMGNEGLKWETLSRNDQVLNNSIARADGLYNTDCSLAP-----566
QY 1011 YLPWPLNDRPNKCPKGGLAAYSTSVNLTSGQVLSRPMAYHKPLKNSQDYTEALRAAR 1070
Db 567 -LLVFLNDHK-----AETLDRAVAQVAFAAENDKDLQFQLAAGNAGI 609
QY 1071 ELAANITADLRKVPCTDPAFEVFPYITINVEOYLITLPEGLFMLSCLVPTP---AV 1126
Db 610 EAAATN-----EVKQSELIILVLVICV-----AAMCMI--TFRSPAAT 646
QY 1127 SCLLGLDLRGLLNLISIVMLVDTVGFMAWDISYNAVSLINILVSAGMSVEFVSHI- 1185
Db 647 LCIVPLILTSVLGNAL-----MAALGIGVKVATLPVLTALGVGIGVDYGIYI 694
QY 1186 TR--SFAISTKPTWLERAKEATI-SMGSV-FAGVAMTNLPGILVLGLAKAQLIOI-PFF 1240
Db 695 TRLESFLRMGLP--LQEAYYETLRSTGKAVLTGLC-----LAIGVATWTFSAIKFOA 745
QY 1241 RLNLLIT---LLGLLHGLVFLPVILSYGVPDVPNALALEOKRA 1280
Db 746 DMGLMLTFMLLNMWFGALWLLPALARFL---INPA-KLQARKA 784

RESULT 81

US-09-328-352-8064
; Sequence 8064, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8064
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8064

Query Match 1.8%; Score 123.5; DB 2; Length 1065;
Best Local Similarity 17.5%; Pred. No. 0.094;
Matches 200; Conservative 178; Mismatches 409; Indels 359; Gaps 54;
QY 344 WGTW-VASWPLTILV-LSVIPV-----ALAAGLVFTTELTTDPVELMSAPNSQAR 391
Db 43 WGTWAVKNTFVDALPDLSDQVVIIRTNFSGQAPQIVENQVYPLATT---TMLSVPGVKTV 99
QY 392 SEKAF-----HDQHPGFFRTNQVI-----LTAPNRS-----XYDS 424
Db 100 RGYSPFGDSFVVYIFDEHTDLYWARSVLEYLNQIGKMPANAKSSLGDPATGVGVVY 159
QY 425 LLLGPKNSGILDLDLLELLEQLRHLQVWSPQAORNISLQDICYAPLNPDTNSLYD 484
Db 160 ALVDP---TGQHDL-----SELRSIQDFWFLKYELK-TLPNVAEVATIGGMVKQY 205
QY 485 CCIN-----SLLOYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGKTALAL 540
Db 206 VVLDPKMATLGVTONNVIEAIQKANOETGGSVLEMAETEYVMVRASGYLTKLDDFRQIPL 265

Qy	541	SCMADYGAPVPFFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWERAFLEEM	600
Db	266	RTNS--PGVPV-----SLGD-----VATQLG-----PEMRAG-----ISEL	295
Qy	601	RAFORMAGMF-----QVTFTAERSLEDEINRTT--AEDLPFIATISYIV-----	642
Db	296	NGOBTGVGVVILRAGKNARETITAVKAKLAELQOSLPKGQVQVVPYDRSOLIDRAVENTL	355
Qy	643	-----IFLYTISLALGSYSSRSRVWVDSKATLGLGGVAVILGAVMAAMPFSYIGRSSL	696
Db	356	SHKLIEEPIVVALVCGFLWHLRGSAMVAIVSLPLG-----ILSAPLVMHYQGLNANI	407
Qy	697	VILQVVPVLVLVSGA-----DNPIFVLETVQRLPRRPGEBPREVH-----IGRALGR	742
Db	408	MSLGGI--ATAIGAMWDAALVMVENAHKHTETQW--EHPQVLEAQVRWNIITRSASE	462
Qy	743	VAPSMLLC--SLSEAI CPFPGA-----LTPMPAVRTFAL--TSGLAVIDFLQWSAF	791
Db	463	VGPTLFPCLLIITLISFPIFTLOQBGLRFLPFAFTKYTMAAAGUSITLIPVL-MGYW	521
Qy	792	VALLSLOSROEASRLDVCCCKPQELPPQCGEGLLIGFPQKAYAPFLHWTGRGVLL	851
Db	522	I-----RGKL-----PSEQRNPLNRFLIKIYSPMLDKVLAHPKTIL	557
Qy	852	L-FLALFGVSUYSMCHIS-----VGLDQ-----ELALP-----KDSYLLDYFLFLNRYFE	895
Db	558	LGALLIFLISLFPFLTRLGGEPFLPNWDEGLLYMPSALPGLSAAKASELQ-----Q	608
Qy	896	VGAPVYFTTLYGYNFSSEAGHNAICSSAGCNPFSTQKIQVATPEPQSYLAIPASSWVD	955
Db	609	TDRMIKTVPVATVFGKAGRAESATDSAPLEMFTT--IQPK-----PRSEWRS	655
Qy	956	DFIDWLPSSCCRLYISGNPKD-----FCPSTVNSLNCCLKNCMSITMG-----	999
Db	656	G-----MTPDKLIK-----ELDKAVQVPGLTNIWVPPPIRNRIDMLATGVKSPIGIKVSAN	705
Qy	1000	-----SVRPSVEQPHKYP-----WFLNDR-----	1019
Db	706	DLQIDIRVAQIQEVAQKMPGVSSALAERLTGGRYVDVDINRMOAARYGLNKDVQOQIVS	765
Qy	1020	-----PNIKCPKGGLAAYSTSVN-----LTSQGVQLASRFMAYHK--	1054
Db	766	SAIGGENIGETVEGLARYPINVRYPREIRDSLEALRNLPILTESGQOIVLSSVANIQTID	825
Qy	1055	--PLKNSQD-----YTEALRAARELAANITADLRKVPQTDPAFYPPYTTITNVFYEQY	1105
Db	826	GPPMLKSENARPPSGWVYVDV--QGRDL-ASVVDQLKQAIDQ-----KVKRSSAMSISYSGQ	878
Qy	1106	LTIPEGLFMLSCLVPTFAVSCLLIGH---DLRSGILNLLSIVMILVDTVGFMAWDIS	1162
Db	879	FERFERANARLKVPIPTIMLIFLLLYLIFRQVQDAVLIMATLPFALIGIWMYLSDYH	938
Qy	1163	YNVAVSLNLVSAGMWSVEF-----VSHITRSFAISTKPTWLERAKATISNGSA	1211
Db	939	FSVAIVGFIALAGVAAEFGVMVLFYKQAEHAQQSLSSST-----ASLTFEQQLNEA	991
Qy	1212	VFAGVAMTNLPGILVLGLAKAQLTIQIFFRNLNLIIT-----LLGLLHGLVFLFPI	1261
Db	992	IHTGAVLRVPRKAMTVAVILAGLPI-----LLGTGTGSELSMRIALPWPVGGMISAPLL	1045
Qy	1262	LSVYGP	1267
Db	1046	SMEVIP	1051

RESULT 82
US-09-252-991A-21000
; Sequence 21000, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

```

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21000
; LENGTH: 1033
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21000

```

Query Match	1.84;	Score 123;	DB 2;	Length 1033;
Best Local Similarity	26.24;	Pred. No. 0.1;		
Matches	70;	Conservative	39;	Mismatches 110; Indels 48; Gaps 11;
Qy	1017	NDRPNKCPGGGLAAVSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAANI	1076	
Db	276	NDPPAPRRMRPMGDAIGLAVAMKPGGDIL-----VLGKALE-----TEFARLQQSLPAGL	325	
Qy	1077	TADLRKVPGGDPAFEVFFPYITWVFYEQYLTIPEGLFMLSCLVETFAVSCLLGLDLR	1136	
Db	326	--ELRKVSDQDPAVR-----TGV--GEFIRLVAEALVIVLL-----VSPFSLG--LR	366	
Qy	1137	SGLINLISIVMILVDTVGFMALDISVNAVLLNLVSAVGMSVEFVSHITRSFAISTKPT	1196	
Db	367	TGLVVALSIPVLAMTFANMHHYEGILHKISIGALVIALGLLVDDAIIVEMAVVMEQG	426	
Qy	1197	WLRAKEATISMGSAPVAGVAMTNPLCPILVLGLAKAQ-----LIQIFFPRLNLIIT	1247	
Db	427	Y-DRLKAAAPAWTSTAPFPMLTGLTLTAAGPLPIATAQSGTGEVTRSLFQVVITL-----	480	
Qy	1248	LIGLLHGLVLPVLISYVGDVDPNPALA	1274	
Db	481	VWSWFAAVVEVP-----YLGAKLIPDLIA	503	

RESULT 83
 US-09-489-039A-13014
 ; Sequence 13014, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 13014
 ; LENGTH: 1049
 ; TYPE: PrT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-13014

	Query Match	1.8%; Score 123; DB 2; Length 1049;
	Best Local Similarity	18.2%; Pred.No. 0.1;
	Matches	146; Conservative 107; Mismatches 244; Indels 304; Gaps 34;
Qy	624 EIRNTRAEIDPIPATSVIVIFLISALGSGYSSWRVNDKATL--GLGGVAVLG--A	679
	: :	:
Dq	340 EVVKTLVE-----AILLVLMYLPLNF-----RATLIETAPVPVLLGTFA	382
	: :	:
Qy	680 VMAAMGFPSYLGRSSILVILQVVPFLVLVG--ADNIFIIVLEVRQLPRPGEPREHVIG	737
	: :	:
Dq	383 VLAAFGF-----SINTLTMGF-VMLATGLLVDDDAIVVENVERVMAEGLPPKEATR	433
	: :	:
Qy	738 RALGR-----VAPSMILCSISEAICFFPGLALTWPAVRTALTSLGLAVILDFELQMSAFV	792
	: :	:

Db 434 KSMQIQALVGIANVLSAVFIPMAFFGGSGTGA1--YRQFSITIVSAMLSVLVALILTP 491
Qy 793 ALLSLDSKROBASRLDVCCCVKQELPPPGGEGELGFFQKAYAPFLHWHI----- 844
Db 492 ALCA-----TMLKPIQKSGHGATTG-FFGWFNRMPDKSTHHTYDTSVGNILR 536
Qy 845 TRGVULLLFLALF-----GVSLSMCHISVGLDOE----- 874
Db 537 STGRYLVLIIIVGMAWLFVRLPSSFLPDEDOGVFL-SMAQLPAGATQERTQKVLDEMT 595
Qy 875 ---LALPKDS---YLLDYFLFLNR----- 892
Db 596 NYLTKEKDNVESFVAFNGFAGRGQNTGIAFVSLKDWSORPGEENKVEAITARAMGYF 655
Qy 893 ---YFEVGAPVYFVTTLGVNFSSEAGMNAICSSACNNFSFTQ----- 932
Db 656 SQIKDAMVFAFNLPAIVELGTATGDFPE-----LIDQGLGHEKLTQARNQLFGMVAQ 708
Qy 933 ---KIQYATEFPEQSYLAIPAS-----SWYDDFDLWLTTP 963
Db 709 HPDLVTGVRPNGLETPQFKIDIDQEKALGVLSISDINTTLGAAMGSGSYNDFID---R 765
Qy 964 SSCRRLYISGPNKDFCPSVTNSLNCNKCMSTMGSVRPSVEQPHKYLPLFLNDRPNIK 1023
Db 766 GRVKVYIMSEAKYRMLPEDIGK-----WYVRGSDGQM 798
Qy 1024 CPKGLAAYST-----SVNITSQDVLSAFRMAYHKPLKNSQDYTEALRAARELANIT 1077
Db 799 VP---FSAFSTSRWEYSGPRLERYNGLPSLEILGOAAPGKST---GEAMALMEELAGKLP 852
Qy 1078 ADLRKVPGTDAFEVPPPTIITNVFEQVLT-----ILPEGLFMLSCLG----- 1120
Db 853 SGI-----GYDWTGMSYQERLSQGNQAPALYAIISLVVFLCLALAYESWSIPP 899
Qy 1121 ---VPTFAVSCLLGLDLRSLGLNLSIVMILVDVTGFMALWDISYNAVSLNL----- 1171
Db 900 SVMLVPLGVGVGALLAA--TFRGLTNDVYFQVGLLTTIGLSA-----KNAILIVEFAKDL 952
Qy 1172 ---VSAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMT 1219
Db 953 MEKEKGLEATLEAVRMRLRPILMTSLAFLGVMP-----LVITSSG-----AGSGAQ 1000
Qy 1220 NLPGLVL-GLAKAQLIQIFF 1239
Db 1001 NAVGTGVMGGMVTTATILAIF 1021

RESULT 84

US-09-252-991A-30591
; Sequence 30591, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30591
; LENGTH: 1052
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30591

Query Match 1.8%; Score 122.5; DB 2; Length 1052;
Best Local Similarity 17.6%; Pred. No. 0.11;
Matches 234; Conservative 166; Mismatches 400; Indels 531; Gaps 57;

Qy 134 SLFTNTR---VAQLGAGQLPAV-----VAYEAFYQHSFAEQSYDCSRVRVPAAATLAVG 186
Db 25 ALFISLAGLIVISKLPVAPQPNVAPPQITITATYP-----GASAKVLVD 68
Qy 187 TMCQVYSGALCNAQRWLNFQO-DTNGGLAPLDITFHLLEPG----- 226
Db 69 SVTSVLSESLNGAKGLLYFESTNNSNGTAETVWTF--EPGTDPLAQDVQVNRLLKKAEB 125
Qy 227 ---QAV-GSGIQ-----PLNEGVARCNEQDDVATCSCODCAASCATAR 268
Db 126 RMPQAVLTQGLQVBEQTSAGFLLIYALSYKEGAQRSDTTALGDYAAARNINNELRLPGVGK 185
Qy 269 PQALDSFYLGQMPGSLVLIILCSVFANVTILLVGRVAPARDKSKMVDPKK---GTS 324
Db 186 LQFFSSE-----AMRV-----WIDPQLVFGULS 210
Qy 325 LSDKLSTHTLLGQFGQMGWVASWPLTILVLSVIPVVAL--AAGLVFTELTTPVEL 382
Db 211 IDD---VSNAIRQNVQ-----VPAGAFGSAFSGSSAQELTATLAVK 248
Qy 383 WSAFNSQARSEKAFDQHFQFFFTNQVILTAPNRSVRYDSLLGPKNFSGLDLDLL 442
Db 249 GTLDDPQBFQGVILRANQDGLVRLADVARLELGEKESYNISSRLNGTPTVGGAIQLSPGA 308
Qy 443 ELLE---LOBRLRHQLQWSPQAQRNISLODI CYAPLNPNNTSLY-----DCCINSLIQ- 492
Db 309 NAIQTATILVKRLAELSAFPE-----DMQYSV--PYDTSRFVDVAIEKVHTLIEA 358
Qy 493 ---YFQNNRTLL--LTFANOTLMQOTSOVDKWHFLYCANAPLTFKQGTALAL 540
Db 359 MVLVFLGMFLQNVRTLIPSI VVVPVCLLGLTLM-----YLLGFSVNMNTMFGMVLAI 413
Qy 541 SMADYGAPEFPFLAIGGYKDKDSEAEALITWTSNNYPAGDPRLAOKLWEAFLEEM 600
Db 414 GILVD-----DAIVVENVERIMA-EGISPA-----EATVKAM 446
Qy 601 RAFQRMAGMFOVTFABERSLEDEINRTTABDLPI--PATSVIVIFLYISLAL----- 651
Db 447 KOVSGAIVGITLV-----LSAVFLPLAFWAGSVGVIIYQOFSVSLAVSILFS 492
Qy 652 ---GYSYSWSR-----VMVDSKATLGLGGVAV 675
Db 493 GFALITFTPALCATULLKPIPEGHEKRGFFGAFNRGFARVTERYSLLNSKLVARAGPML 552
Qy 676 VLGAVMAAGFPFSLGIRSSLVILQVFPFLVLCADNIFIFVLEYQRLPRPGEPEVH 735
Db 553 VYAGLVAMLYGF-YLRLEPAFVPAEDLGYMWVDV-----QLP--PGASR--- 593
Qy 736 IGRALGRVAPSMLLCSLSEALCFILGALTMPAVRTFA-----LTSGLAVILDFLQ 787
Db 594 ---VTDATGELEERFLKREAVASVFLIS 620
Qy 788 MSAFVALLSLDSKROBASRLDVCCCVKQELPPPGGEGLLLGFFQKAYAPFLHWHITRG 847
Db 621 GFSF-----SQGDNAALAF-----PTFKDWSRG 645
Qy 848 VVLLFLFALFGVLSYSMCHISVGLDQELALPKDSYLL----- 884
Db 646 ---AEQSAABEIAALNEHFALPDGVTMAVSPPPINGLNGSGGFALRLMDR 693
Qy 885 ---DYFL-----FLNRYFE--VGAPVYFV-----TTLGYNFSSEAG 915
Db 694 TGVGREALLQARDTLIGEITQNPFLYAMMEGLAEAPQLRLIIDREKARALGVSFETISG 753
Qy 916 -MNAICSSAGCNGNFSFTOKIQVATEFPPEQSYLAIPASSWVDVDFIDWLTSSCCRLYISGP 974
Db 754 TLSAAFGSEVINDFTNAGRQQRVVIQAEQG-----NRMTPESVLELYV--- 796
Qy 975 NKDKFCPSVTNSLNCNKCMSTTM--GSVRPSVEQPHKYLPLFLNDRPNIKCPKGLAAY 1032
Db 797 ---PNAAGNLVPLSAFVSVKWEQPV-----QLVRY-----NGYPSIRIVGDAAPGF 840

```
QY 1033 STSVNLTSDQVLAARFMAHKLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEV 1092
DB 841 STG-----EAAEMERLAAQAPAGI-----860
QY 1093 FPYTITNVFEOYLTI-LPEGLFMLSICLVPTFAVSCLLGLDLRSLNLLSIVMIL-- 1149
DB 861 -GYEWTLGSLYOEKVSAGQATSLPALILV-----FLLVALYBSWSIPUSMUIVPI 912
QY 1150 --VDTVGFMALMDISYNAVSLINLVSAGVMS-----VEFVSHITRSPAISTKPTWLE- 1199
DB 913 GAIGAVLAVMVSGMSNDVYFKVGLITITIGLSAKNAILIVEF-----AKELWEOG 961
QY 1200 -RAKEATISGSAVAGVAMTNLP---GILVL-----GLAKAQLIQIFFRNLNLLITLL 1249
DB 962 HSLRDAIAEARLRPRITMSMAFILGVIPLALASGAGAASQRAIGTGVIGMLSATEL 1021
QY 1250 GLLHGLVFLPV 1260
DB 1022 GVL-----FVPI 1028

RESULT 85
US-09-902-540-10725
; Sequence 10725, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10725
; LENGTH: 863
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-10725

Query Match 1.7%; Score 120; DB 2; Length 863;
Best Local Similarity 19.9%; Pred. No. 0.14;
Matches 154; Conservative 120; Mismatches 328; Indels 172; Gaps 34;

QY 549 PVFPFLAIGYKGYSEABALIMTFSLNYPAGDPRLAQAKLWEE-AFLEEMRAFQRRM 607
DB 172 PMWDNTEIG--KTKDYLER---LKGQIDAYSA--QPGKVKLVEDYKLMGDGKTIAYGF 222
QY 608 AQMFQVTFTAERSLEDEINRTTAEDLPFATSYVIVFLYISALGSYSWSRWVDSKAT 667
DB 223 TGSYKITVDDSPAIBESLEPVTL-----IALGSIFLITITIFFRKLAP 264
QY 668 LGLGGVAVVLGAVMAAGPFYSILGRSSILVQLVPLVLSVGADNIFVLEYORLPER 727
DB 265 TFLVLVIGTVGVIY-TLG-FTYATVGLNMTISILGILLGIDYGHFVR-TRLELG 321
QY 728 PQGEVPHIGRALGRVAPSMILCSISEATCFPLGALTTPMPAVRTFALTSGLAVIDFLLIQ 787
DB 322 ACKPYDVAIRDVAMNAGRAAUAUVTAGSPFVLVMSBEPFGSQFGLAGMTLILGLTL 381
QY 788 MSAPVALLSLDSKREASRLDVCCCKVQDEL-----PPPGQEGELLGPFQKAYAPFLIH 842
DB 382 FCWSAAILAL-----AGRINP--ELPQKLIGVMKPP--TNSATTGKELRIPKPMVL 430
QY 843 WITRGVLLLLFLALFGVSLYSMSCHISV---GLDQELALPKDSYLLDYFLFNRYFEVGAP 899
DB 431 GLSTAIVALI-----CAAAPWAGSE--PPKGVELGFFERLK-----466
QY 900 VYFVTLTGVNFSEEA-----GMAACSSAGCN---NFSFTQKIQYATEPPE-----Q 943
```

```
DB 467 -YGVS---FNYNTRALIPDGMSSVLLQDEINERFNISDPMATYTTKDLDEAGVYRELQ 522
QY 944 SYLALPASSWDDFDIMLTPSSCC-----RLYISGPNKDKFCPSTVNSLN 988
DB 523 NAHKPSIDQVSVISFTFVPVPAETAANAANAKVLSEWKAEMQOLEDEGFSVALPPEMANAD 582
QY 989 CLKNCMSITMGSVR--PSVEQFHKYLPMFLNDRPNIKCPKGGLAAY-STSVNLTSDQVQL 1045
DB 583 FFKVKLDAKFPDVGHPA---NYTAQFEN-LPSAKPENHGVLTYYIASVDL-MDQ-- 633
QY 1046 ASRFWAYHKPLKNSQDYTEALRAA-----RELAANITADLRKVPCTDPAFEVFP 1094
DB 634 -----KMLKFS-DETKVIAAATPGKFDQDAMDPKAPTVEKEPRAAGAT-----QLYA 680
QY 1095 YTITNVFEOYLTIPEGLFMLSICLVPTFAVSCLLGLDLRSLNLLSIVMILVDT-- 1152
DB 681 RLARIVLWDGKVTVLTAIWIAMHP-----LDFRNVKALASVPLGVGVAM 728
QY 1153 -VGFMALMDISYNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMGSA 1211
DB 729 MLGIMALTGLRLNFMNIIILPILGFGVSHGLYLLHRFLEGTSPVALRSVGAAVA-SST 787
QY 1212 VPAGVAMTNLPGLIVLGLAKAQLIQIFFRNLNLLITLGLHGLVPLPVLISV 1265
DB 788 LTAVVAFAALLAAAHNGL-----RSMGLVACIGLITTLVVSFTVLAHV 830

RESULT 86
US-09-902-540-14015
; Sequence 14015, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14015
; LENGTH: 735
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-14015

Query Match 1.7%; Score 118.5; DB 2; Length 735;
Best Local Similarity 27.5%; Pred. No. 0.15;
Matches 53; Conservative 32; Mismatches 59; Indels 49; Gaps 10;

QY 1136 RSGLLN-LLSIVMILVDT-----VGFMALMDISYNAVSLINLYSA----- 1174
DB 150 RAGTLSALIAVGLIILATERVAALAVGIPVMGVVLIFFAAQAQLAIGHLNIYVGLVAL 209
QY 1175 VQMSVEFVSHITRSPAISTKPTWLER-----AKEATISGSAVAGV---AMTNLPGLIVL 1227
DB 210 IGLGIEYGVHLCMYR-----WEERRAHAPAREALVTAVRGTFSGAVTSVNTAAAPFVL 262
QY 1228 GLAKAQLIQIFFRNLNLLITLGLHGL-VFLPVLISY-VGPDVNPALALEOKRAEEAVA 1285
DB 263 LLAQOFAQNQP-----GLLAGLGVLLAVLATVAMGPSL-LALAERLRPARVDVA 310
QY 1286 AMVASCNHPSR 1298
DB 311 AEATAPOVSQPER 323

RESULT 87
US-09-252-991A-20611
; Sequence 20611, Application US/09252991A
```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20611
; LENGTH: 1033
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20611

Query Match 1.7%; Score 118.5; DB 2; Length 1033;
Best Local Similarity 21.0%; Pred. No. 0.27;
Matches 163; Conservative 111; Mismatches 296; Indels 207; Gaps 40;

QY 623 DEINRTTAEDLPFATSVIVIFLYISLALGSYSSMSRVMDSKATLGLGSAVVLG--AV 680
DB 335 DEVVTLVEAL-----LIVLVIVLCLGSL-----RSVLIPVATIPLS-----MLGAAL 379
QY 681 MAAMGFFSYGIRSSILVQVFFVLVSVGADNIFIVLEYQRLPRRPGEPREHIGRAL 740
DB 380 MLAFGFSVNL-----LTLAMV--LAIGLVDDAIVVW---ENVHRHIEEGKSPVAAALI 429
QY 741 G--RVAPSMLLCSLEACF-----FLGALTMPAVRTALTSGLAVID--FLLQMSAFV 792
DB 430 GAREVAGPVIAVTITLAAVYTPIGLMGGLTG--ALFPREALTAGAVIVSGVVVALTSPVM 488
QY 793 ALLSLDSKROEASRLDVCCKVPQBLPPPGQEGLLGFQKAYAPFL-----LHWITRG 847
DB 489 SELLQAQNEGRM-----GRAEWFPGGLTRYQVLEFSLGRHWLTCG 533
QY 848 VVLLFLPLAFGVLSYSMSCHISVGLDQELALPKDSYLLDYFLPLNRYFEVGAPEVYFTILG 907
DB 534 LALLVCISL--PLLYSM-----PKRELAPTED----- 558
QY 908 YNFSSEAGHNAICSSAGCNPFSTQKIQVATEFP--EQSYLAIP--ASSWDDFDIMLT 962
DB 559 -----QAAVLTALAKOAHN-----LDYVELPARKLDQVYTSIPETVSTWIIINGTDG-P 606
QY 963 PSSCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFH-KYLPWFNLDRP- 1020
DB 607 AASFGINLAAMEKREERDASAIQSE--LQKVGDEVS--SIFAFQALALPGSTGGLPV 661
QY 1021 --NICKPKGGLAAYSTSVNLTSDGQ-----VLASRPMYHKPL-----KNS----- 1059
DB 662 QMVLRSQDPYVLYRTMBEIKOKARQSGLFVVVDSDLDYNNPVQVDRAKANSIGIRM 721
QY 1060 QDYTEALRAARELANITADLRKVPCTDPAFVFPVYTIITNVEQVLTILPGL---FML 1116
DB 722 QDIGESLAV--LVGENYNNRFGMEGR--SYDVIPOSRL-----DQRF--POLARARQFVR 770
QY 1117 SL--CLVPTFAVSCILLGLDRLSGLNLLSIVMILVDVTGFMALMDISYNAVSLNLSA 1174
DB 771 TDQGNLVPLSTVVRVALQVEPNK-----LIQPDQQAATLQAIAPAGVSM----- 815
QY 1175 VGMSEVFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTNLPQILVLGLAKAQL 1234
DB 816 -GQAVAFLDVARGLPAGFSDHWQSDRSQYTOEGNTLVEAFLA-----ALVVIYLVLAQ 869
QY 1235 IQIFFFRMLLITLLGLLHGLVFLPVLSVVGDNVNPALAE----- 1276
DB 870 YBSLADPLIITVPLUSICG--ALLPLAAGVATNNTYIQGLVTLIGLSKHGILMVEFAN 928
QY 1277 ----QKRABEAAVAVMVASCPNHPSPRVSTADNIYNHSPFGSI-----KGAGAINSP 1324

Db 929 ELQLHERLDRRAAILRAAQIRLRPLVMTTAAWVF-----GLVPLLFSAGAGAAASRF 979

RESULT 88
US-10-104-047-2923
; Sequence 2923, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2923
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2923

Query Match 1.7%; Score 115.5; DB 2; Length 914;
Best Local Similarity 18.8%; Pred. No. 0.41;
Matches 208; Conservative 140; Mismatches 381; Indels 375; Gaps 55;

QY 55 VCSLSNTPARKITG-----DHLIL-----LQKICPLRYTGP-NTQACCSAKQLVS 98
DB 19 VICLPRPP--KVLGYRLEPPHLLTLACTLEGMYNIEELERDINGTLLSQATCELCD--G 74
QY 99 LEASLSITKALLTRCPACSDNFVNLHCHNTCS-PN---QSLPFINVT-----RV AOLGAG 148
DB 75 NENSMVNVNLDRCVRCCEPTFVNTSRSCASEPNILTGGLCFSTGNFPLRISAARYG 134
QY 149 QLPVAVYEAIFYQHSFABQSYDSCSRVRVPAATAALAVGTCM---GVYGSALCNAORMLN 204
DB 135 EVGASLTSEWFAKYL---QSSAAACWVYANLTSCQALGSMCMVMNNSYDFATFDACGLFQ 191
QY 205 FQDGTGNGLAELDITFHLL-----PCQAVSGTQPLNEGVARCNESQGDVATCSCQD 258
DB 192 FIFENTAGLS---TVHSISFWRQNLPMWLFYGDQL-----GLA----- 225
QY 259 CAASCAPALARPOALDSTFLGQMPGSLVLIILCSVPFAVVITLLVGFVRVAPARKSKMVD 318
DB 226 -----PQVLSST-----SLPTNPSF----- 240
QY 319 PKGTSLSDKLSF--STHTLLGQFPQGGTGWVSWPLTILVLSVIPVALAAGLVFTLT 376
DB 241 --KGBDQNTKLKFAAASYDIRGNFLK-WQT-----LEGGV----- 272
QY 377 TDPVELMSAPNSQARSEKAFHDQHPGPFRTN-----QVILTAPNRSSRYRDSLLGPK 430
DB 273 -----LQCPPTETRLNAA---SEFTTYQQNCEPIPSKILIDFPPTIFY----- 314
QY 431 NFSGILDLDLLELELELERLRLQVMSPEAORNISLQDICVAPLNPNONTSLDYCCINSL 490
DB 315 -----DVLEYTDENQHQYILAV---PVLNINLQH-----NKL 344
QY 491 LQYFQNNRLLLLLTANQTLMGQTSQVDWKDHPLYCANAPLTPKDGTAALASCMA---DYG 547
DB 345 FVNQDSNSGKMLLTRRIFLV---DAVSGRENDL--GTQPRVIRVATQISLSVHLVPTNTIN 399
QY 548 BPVFFFLAIGGYKGYDYSEAE-----ALIMTFSLNNYPAG--DPLRAQAKLWEERAPLEEM 600
DB 400 GNIYFPLITIAVSDIDIKDANSQSVKVSFVSVTYEMDGEAHVQTDIALGVGLGLAVLASL 459
QY 601 --RAPFORMAGFQVFTTAERSLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSSWS 658
DB 460 LKTAGKRRIGSPMIDL-----QTVVKELVYVAGDLANVFFIITVGTGLY--W- 505
QY 659 RVWDSKATLGLGGVAVVLGVMAAGFPFSYLGIRSSLVILQVVPFLVTVSGADNIFIV 718

Db 506 --LIPFKAO---KSVSVLPMPOBERFVTVGCAFALAKALQFLHKLISQITID---VFF 557
Qy 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSBAICFFFLGALTMPFAVRTP----- 772
Db 558 IDWER-----PK-----GKVLKAV-----EGEGVRSATVPVSIWRTYFVANEM 596
Qy 773 -----ALTSGLAVILDFFLOWSAFVALLSLS-----KQEBASRLDVCCKVPQBL 818
Db 597 NBIQTVRKINSLFQVLTVL--PFLEVVGFKNALMDSSSLSRNPSPYIAPYSCI----- 649
Qy 819 PPPQGEGLLGFQKAYAPFLHWTITRGVWLLPLALF-----GVSLSXSMCHI 867
Db 650 -----LRYAVSAALMLAIGLIQVVFAYFIEPIEDKIQFVDLCSMSNI 694
Qy 868 SVGLDQELALPKDSYLLDYFLNRYFVGVAPY-----FVTTLGYNFSSEAGMNAICSS 922
Db 695 SV-----PFLSHKCF--GYIIGHRSVGHADTNMBEEMNNLAKREA--ENLCSQ 738
Qy 923 AGC--NNFSFTQKIQYATFPPQSYLAIPASSWVDDFDLWLPSPSCCRILYISGPNKDKFC 980
Db 739 RGLVPNTDQGTPEIAISNQ--RQYDRIHET-----LIRKNGPARLLSS----- 782
Qy 981 PSTVNSLNCCKMCSITMGSRVPSVEQPHKYLFPWFLNDRPNIKCPKGGLAAYSTSVNLTS 1040
Db 783 ASTFQOSIKAYHMMNKFLSP---IDHVHKMDYFIKDKLLLE----- 822
Qy 1041 DQQVLASRFMAYHKPLKNSQDYTE 1064
Db 823 --RILGMEFM---EPMEKSIFYND 841

RESULT 89

US-09-902-540-11066
; Sequence 11066, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11066
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11066

Query Match 1.7%; Score 115; DB 2; Length 651;
Best Local Similarity 19.8%; Pred. No. 0.25;
Matches 129; Conservative 90; Mismatches 228; Indels 206; Gaps 32;
Qy 821 PGQEGGLLGFQKAYAPFLHWTITRGVLL-----LFLAFGVSLYSMSCHISVGLDQEL 875
Db 35 PAQME-----GYFTKMVNNML--FTNGIKNIETKNSQGLTLMKLSFYE-----GTDMGQ 82
Qy 876 ALPKDSYLLDYFLNLR-----YFEGAPVYFVTTLGYNFSSEAGMNAICSSAGCNFFSFTQ 932
Db 83 AVAEINAL-----SNRSQVFLPPGAPPPFI--IRFDASSHPVGHVLPSES-----KTTN 130
Qy 933 KIQYATFPPQSYL--AIPASSWVDDP-----IDWLTP-----SSCCRLY 970
Db 131 QLQDIANFTARPFLISIPGLTAPPGGSPRTIEINIEPEKLRVHNLTPQVVEAIARON 190
Qy 971 ISGPNKDKFCPSTVNSLNCCKMCSITMGSRVPSVEQPHKYLFPWFLNDRPNIKCPKGGLA 1030
Db 191 VTAPSGNVHGVGVD-----TYITPTNSTLR-AVEDF-----GNIPLLKGSVA 230

Qy 1031 -AYSTSVNLTSDQGVLASRFMAYHKPLKNSQD--YTEALRAARELAANTADLRKVPGETD 1087
Db 231 NVYVRDVATVXGADIATGY-----ALVNGKRSVYLVNAKAGNASTVHVVVQQLK----- 279
Qy 1088 PAFEVPPYITIN-----VFYE-----QYLTILPEGLPMLSLCLVPTTFPAYSCLLGLDLRS 1137
Db 280 ---ESIPRIQSMLPDDVHISYBFDQSVYVVGALKGLIAGEVIGAVLTGLMWLLFLRDVRS 336
Qy 1138 GLLNLLSLVIMILVDTVGFPMALWDISYNAVSLNLNLSAVGMSVBFVSHITRSPFAISPKPTW 1197
Db 337 AIIVITTIPIAISGLVFLKFGQTINIMTLAGLALAGVILD-----ESTVTIENIHQH 391
Qy 1198 LERAKAETISMGSA-----VFA--GVAMTNLPGLILVLGLAKA---OLIQ 1236
Db 392 LQRGKTAVAVMWDACLEIAFASKLLILLCILSVFAPALTMGGIPGALFRPLALAIGCMVV 451
Qy 1237 IFFFLNLLITLL-----GLL-----HGLVP 1257
Db 452 SFLLSQSLVPVNMNMLKQHPGPGGASPSGRRVRVIEGMLPHRKLILLVAGVGLIV 511
Qy 1258 LPVI--LSYVGPDPNPAI-----ALEQKR-----AEEAVAA---VM 1288
Db 512 LAVVSLQRIQKDVLRVNSSQLQLRLRAAEGTRIEKTEQVHVHQMGLVEEVVAGKVRIS 571
Qy 1289 VASCPNHPSRVSTANIYVNH-----SFEKSIKAGAISNPLPNNGRQ 1331
Db 572 SAYVQHPSPFAISP-IYLYNAGPHEALLQVAFEGVGVDIDALKERIRQVRRE 623

RESULT 90

US-09-902-540-12414
; Sequence 12414, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12414
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12414

Query Match 1.7%; Score 115; DB 2; Length 810;
Best Local Similarity 20.0%; Pred. No. 0.37;
Matches 169; Conservative 102; Mismatches 301; Indels 274; Gaps 37;
Qy 551 PFPLAIGYKGYKDYSEAEA-----LIMITFSLNNYPAGDPR 585
Db 124 PATLGVSPVIGRAFTAEAVQGRBVRVVLTDATWRVHPARDAQVLGQTMLD-----GEFY 179
Qy 586 LAQAKL-----WEBAFL-----EEMRAPORMAGMPQVTTFAERSLEDEINRTTARD 632
Db 180 TVVGVLPAGVAYPANAEMVVPVPTETQREQDQRGARFLSVVGRLLKPGVGTQDAART---D 236
Qy 633 LPPIFATSYIVFLYISLALGS--YS-----SWSRVWVDSKATLGLGVAVVILGAVMAAGPF 687
Db 237 LARVA-----LEMEAVPSRKYRKVGWSFSVTSIDKV---VGNVRGTLMILLGAVGP 286
Qy 688 SYLGRSSILVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRAL-----GRV 743
Db 287 LLAACSSVANLL-----LARMARGREVSIRALGASRGEL 322
Qy 744 APSMLLCSL-----SBAICFFLGALTMPA-----VNTPALTSGLAVI 781

```
Db 323 VAQFLTESLVLSLGGVGLMLLALWGTDALVALVGEALPRASQVRLDARPLVFTMGVSL 382
Qy 782 LDFLQMSAFVALLSLDSKROEASRLDVCCCVKQELPPPGQEGELLGFFQKAYAPELL 841
Db 383 TGVLPGLGPAL-----HGREDLISVAMREGSRGTEGHSG-----RMRAGLV 425
Qy 842 HWITGVVLLFLALFGVSLYSMCHISVGLDQ-----LALPKDSYLLDYFLFLNRYFE 895
Db 426 GOVAVALLVGLAGLFTKSFALRAVDAGFTPEGVLTKGLVLP-----ARYPD 474
Qy 896 VGAPYVFTVTLGYNFSSBAGMNAI-----CSSAGCNFSTFKIQIYATEFPEQSYLAIPA 950
Db 475 AAKHVAFORELLGRLOSLPGVEAVGVNLLPLGGRSDSF--DIEGRPOADPEVWPAVEF 532
Qy 951 SSWVDDFDLWT-----PSSCCRILYISGPNKDKFCPTVNSLCLK--- 991
Db 533 RSVSPGYLRTLRVTPREGHLLGEGDTEGPDAPQVQVINKTFADLYWPGQDALGQRLKLHW 592
Qy 992 -NCMSITWGSVPSPVEQPHKYLFWELND--RENIKCPKGLAAYSTSVNLTSDGOVLASR 1048
Db 593 DTAQWTTVVGVIVDDVRE-----WGLDTPARP-----AAYTPAAKVPTFFALAVR 637
Qy 1049 FWAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPVEVPFYTTINVFYEQYLT 1108
Db 638 AKSGSP-----EALRTA-----IEALRAVDGNLALYAVAPLT-----RL 673
Qy 1109 LPEGFLMLSLCVPTFAVSCLLGLDLRSLGNLLSIVMILVDTVGFMAWMDISYNAVSL 1168
Db 674 VDES-----ISSRQVSALLMGLFAGTALL-----LAA 700
Qy 1169 INLVSAVGMVSFVSHITSPFAISTKPTWLERAKEATISMGSAFVAGVAMTNLPGILVLG 1228
Db 701 LGISGVIGYS--VAQRTREMG I -RMALGASRVLVTLVGO-----GLRLTALGVVVLGLG 752
Qy 1229 LA--RAQLTQIF-----FFRLNLIITLLGLLGLVFLPLVILSYVGPDPVNPALAL 1275
Db 753 LSLGLARLLDAMLGYVAAVDGWTAGVALLGTGVAVL--AAMP--ARRATRVDPALAL 807
Qy 1276 EOKRAE 1281
Db 808 ---RAE 810

RESULT 91
US-08-311-731A-363
; Sequence 363, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
```

```
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 363:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
; US-08-311-731A-363

Query Match 1.7%; Score 114; DB 2; Length 976;
Best Local Similarity 18.3%; Pred. No. 0.65;
Matches 151; Conservative 112; Mismatches 255; Indels 306; Gaps 37;

Qy 175 VRVPAATLAVGTMCVGVYGSALCNAQRWLN-----FQDGTGNGLAPLDITPHLIE 224
Db 273 VNLTSLSAIAAGTDYGFITGRYQEARQANENKEAFTMYRG-----TFHVI- 320
Qy 225 PQQAVGSGIQPLNEGVARCNESQGDVATCSCQDCQCAASCPAIPARQALDSTFYLGOMPS 284
Db 321 ---LGSGI-----TIS---GATECLSFARMYPYQF---LG-VP 350
Qy 285 LVLIILCSFVAVTILLVGRVAPARPKSKMDPKKTSLSDKLSFSTHTLLGQFFQGW 344
Db 351 VGMLF-----AVALTILGPAVLTVGSRFLGFEPRKLIK-----RGM 388
Qy 345 ---GTWVASWPLTILVSVIPVVALAAGLVFTELTPDVELWSAPNS-QARSEKAFHDOH 400
Db 389 RRIQVVVRVPLPILITTC--AIAMVGILLALPGYTNKYKDRAYLPASIPANQGFAAADR 446
Qy 401 FGPPERTNQVIL-----LQYFQNNRTLL-----LTA 505
Db 447 F-PQARMKEPILMTESDHDMENPADFLIDKLARGIFRVPGISRVAQITRPDGTAMDHTS 505
Qy 420 Y-----RYDSLILGPKNFS-GILDDLLELLE-LQERLRHLQW 457
Db 506 IPPQISMQVAGQVQTMKYQKDRMNDLLRQAEINMAETIASMRMHQMLMALLTHTHILND 565
Qy 458 SPEAORNIS-LQD-----ICVAPLNPNSTLYDCC----- 486
Db 566 IVEMQRTTSKLDEITANEDDFWRPIRSYFYMERHCFNIPICWS-----FRSIFDALDGDV 620
Qy 487 ---INS-----LQYFQNNRTLL-----LTA 505
Db 621 QIDERLNSIVGDIKNMDLLMPQMLEQFPFMIESMESMRTIMLTHTSTMSGIFDQNMELSD 680
Qy 506 NOTLMGQTSQVDWKDHFLYCANAPLTFKQGTALALASCHADYCAPVPPFLAIGSYKGDYS 565
Db 681 NANTMGKAPDTAKNDDSFYL--PPEVEKN-----TDFKRAMKSPFUSSDGH----- 723
Qy 566 EAEALIMTFSLNYPAGDP-----RLAQKLWEEAFLEEMRAFORRMAGMQV 613
Db 724 AARFIL-----HRGDPASVAGTASINAIHTAAEEALKGTPLEDTKIYLAGTAAVFK- 775
Qy 614 TFTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWRSVMVDSKATLGLGV 673
Db 776 -----DIDEGANWDLVIAGISSLCILFIILII-----TRAFV--AAAVIVGT 817
Qy 674 AVVILGAV--MAAMGFPSYILGIRSSVLVQVFPFLVSLVGADNIFIFVLEYQRLPRRPGEP 731
Db 818 ALSLGASFGSLVLLWQHILGIELHVLVLAWSIVLVLAGVSDYNLLVLSRFK----- 868
Qy 732 REVHIG-----RALGRVAPSMLLCSLSSEACIFFLGLALTFMPAVRTFALTSGLAVIDFL 785
Db 869 QEIQAGLKTGIIRSMGTGKVTNAGLVFAFTMASVMVSDLRVIGQVGTIGLGLLFDTL 928
Qy 786 L-----QMSAFVALLSLDSR-----QEARLSDVCCCKFQELPP 820
```



```
Db 929 IYRSFMPMSIAALLG-----RWFHWPOQGRTRPLLTVAAPVGLPP 968

RESULT 92
US-09-252-991A-24417
; Sequence 24417, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24417
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24417

Query Match 1.7%; Score 114; DB 2; Length 1019;
Best Local Similarity 19.8%; Pred. No. 0.7;
Matches 162; Conservative 129; Mismatches 309; Indels 218; Gaps 43;

QY 580 PAGDP-----RLAQAKLWEAFLEEMRAFORRMAGHGFQVFTFAERSLEDEINRTTADPI 635
Db 289 PSANPLDVIKYVRAMPE---LEEQLPNLKVSIAYDATRFIOASI-DEWVKTLGE---- 340
QY 636 FATSIVIVPLXISLALGSSYSSRVVDSKATLGLGVAVLGVAAWAGPFSYLGRSS 695
Db 341 ---AVLIVVVVFLGAP----RSVLIPVTIPLSMGVLP--FWQMGY-----S 383
QY 696 LVILQVWPLVLSVG--ADNIFIVLEYQRLPRPCEPREVHIGRALGRVAPSMLLCSLS 753
Db 384 INLLTLA-MVLAIGLVADDAIVVENIHRHTEEGKPPPEAALEGAREIAVPVWSMTITL 442
QY 754 EAICFELGNLTMPAV--RTFALTSGLAVILDFLOMSAFVALLSDSKQF-----ASR 806
Db 443 AAVYAPIGLTGLTALGFKEFATAGAVIISGIVALTSLPMNCSRLRHEENPNSGLAHR 502
QY 807 LDVCCVXPQELPPPGQGGLLGFPKAYAPFLHMI--TRGVVL---LLFLALFGVSL 861
Db 503 LDLIF-----EGLKQR-YQRA-----LHGTLDTRPVVLVFAVLVLAIPVLL 543
QY 862 YSMCHISVGLDQELALPKDSYL-----LDYFLFLNRY---PE---VGAPVYF- 902
Db 544 -----MFTKELAPEEDQGIVFLMTNSPQTANLDY---LNRYTAETEGIFRSFPEYYS 593
QY 903 -VTTGYNFPSSBAGNAICSSAGCNNSFTQKIQA-----TEPPEQSYLAIPASSWDD 956
Db 594 AFQINGYN-GVQAGIGGMLLPKWEREKSQEMELLHAVQAKLNEIPGVQIFAFNLPS---- 648
QY 957 FIDMLTPSSCCRLYISGNPKFCFSTVNSLCKNCSITWGSVRSVPQPHKYLPMFL 1016
Db 649 -----LPGTGEGLPFPQVLNTANDYESLIQVAQ-RVKQRAESGKFA--FL 691
QY 1017 N-----DRPNI-----KCPKGGGLAAYSTSVNLT-----DGQVLASRFW- 1050
Db 692 DLDLADFPELVVDIDREKAAQMGVSMQDLGVALASLLGEGEI--NRTIDGRSVKVIQA 749
QY 1051 -----AVHKPLKNSQ-----DYTALRAARELANIATADLRKVPGTDPAPFV 1092
Db 750 VERPYRDNPGWGLSYVKSRLNQALVALSTLIETHERARPRQLNQFQOLNSA-----IISG 804
QY 1093 RPY-----TITNVFVEQYLTILPEGL-----FMLSCLVPTFAVSCLLAGL 1133
Db 805 PPIVSMGEAIEYVQOIAREE-----APRGFANDYAGASQYVQEGSALLVTFGLALAIIFL 860

QY 1134 DL-----RSLGLNLLSIVMLVDTVGFMALWDISYNNVSLINLVSAVG-----MSV 1179
Db 861 VLAQFESFRDPLVIMVTVPLSICGALIPFLFGVSSLNITVQGLVTLGLSKHGILIV 920
QY 1180 RFVSHITRSFALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQLIQIFP 1239
Db 921 EFANQLRHEQGIGRREA-IEQAAAIRLRPVLMTTAAWLVGLVPLILATGAG-----AVSR 974
QY 1240 FRANLLITL---LGLLHGLVFLPVLVSYG-PDVNPAL 1273
Db 975 PDGIVATGMSVGLFTLFLVLPCTIYTLVARDPGV 1012

RESULT 93
US-09-603-208A-226
; Sequence 226, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 226
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-226

Query Match 1.6%; Score 112.5; DB 2; Length 697;
Best Local Similarity 18.8%; Pred. No. 0.5;
Matches 183; Conservative 111; Mismatches 311; Indels 369; Gaps 44;

QY 341 FQGMGTWASWPLTTLVLSVIPWVALAAGL-----VFTLTTDPV--ELWSAPNSQARSK 394
Db 2 FSKGHGFPAYR-----RRIVPLVVIATLALFVIFGTGKGLGRMSQEGWDDPGSSSTAA 55
QY 395 AFHQHGFPPFTNQ-----VILTAPNRSRYDSLLGLPKNPSGILDLLELLELQER 450
Db 56 RIELETFG---RDNDGDVLLFTAPEGTSDDAEVF---SSISGYLD-----GLIE---- 100
```


Query Match
Best Local Similarity 20.5%; Score 112; DB 2; Length 750;
Matches 109; Conservative 73; Mismatches 195; Indels 154; Gaps 26;

QY 308 APARDKSNWDP-----KGTSLSDKLSPSTHTLGGPF-----QGNGTWVASHPLILV 357
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
161 ASAGTRSIADPVGDKDGTPIQAIOPFAGAPKNKEFTTVPVPVG-----KVITSA 212
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
358 LSV--IPVALAAGLVTEL--TTDPVELM-----SAPNSQARSEKAHFHDHGFPPERTN 408
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
213 SNVDVVVVPGGRGVAFIELKDRLPDLNYGTIIPLEATLSLANGASYNESWA-WHRSM 271
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
409 QVILTAP--NRSSRYRYSLLGLPKNFSGILD--LDLLEL-LETQERLRHLQW---SPE 460
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
272 STIRNAVLDERMVELDRL---QEFAALVQQRKDLVSSIQVELARRVQPIQGLKAEPE 328
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
461 AQRNISLODICYAPLNPDNTSLX-----DCCINSLLOYFQNRTLLILLTNQTLMGOTSQ 515
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
329 TSSNLITDDD---EPDPFVTTLQYLASLDVAIEAALQAQAKGCL-----NFIQIGKAAAP 380
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
516 VDMKDHFLYCANAPLTFKDGFTALALSOWADY-GAPVFPELAIGYKGDYSEBALIWP 574
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
381 CDW-----SPKRF-----ARRVMNLYQGAREQDFOKCSTYTTRDDFADLKARELQI 425
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
575 SLNNYPAG-----TFTAERS-----LEDEINRTTAE-- 631
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
426 GTVNPYPAQDYTTSPSRLETTYRRDQYMRALSSVTGPMLDPKTAQPLRWESGDSYSLGN 485
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
603 FORMAGMFQV-----TPAERS-----LEDEINRTTAE-- 631
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
486 DTFCATANYQVAVAINNVKNTPCDGLSPEVTGSPTATGSGALGISANLINAQVNVTQKRAD 545
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
632 -DLPIPATSVIVIFLYISLAGSYSSWS-----RVWDISKATLGLGGVAVVLGVAMMA 684
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
546 IDLDVLDNTVPLVDIHKNLTLGQFNIVSGRKQEKATLVSAQATEFVIWVVPVTIGAKVAGV 605
QY ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
685 GFFSYLGIRSLVILQVVPFLVLSVGAD-----NIFI FVLE 720
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
606 VGLEY-GLSADHKVTGTACTSVTSVGTGHVQPYAQVDGELYASVDLFVIE 655

RESULT 96
US-09-543-681A-8143
; Sequence 8143, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8143
; LENGTH: 1051
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-8143

Query Match
Best Local Similarity 20.1%; Score 112; DB 2; Length 1051;
Matches 125; Conservative 87; Mismatches 193; Indels 216; Gaps 31;

QY 811 CC---VKPQLPPPGQEGELLGFFQKAYAPFLHWITRGVLLLFIA---LFGVSLSYM 864
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
3 CCGESIMSEKIPATGGGPSRL-----FILRPVATTLPFWAILAGIVGYRM 48
QY 865 CHLSVGLDQELALPKDSLDDYFLPLNRYPEVGAPYVFTTL--GVN----- 909
DB ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
49 LPVMS-----ALP-----EVDYPTIQVVTLYPGANPEWMTSAVTAPLE 85

Qy	910	--FSSAGMNAIC--SSAGCN-----NFSFT-----QKIQYA-----TEFP----	941
Db	86	RQFGMSGLKQMSQSSGGASVITLAFQLTPLDVAEQVQAANAATNLLPTDLFPYPI	145
Qy	942	-----EQSYLAIPASSWDDFDMLTPSSCRLYISGPKDKFCPTVNSLNCIKNCM	994
Db	146	YSKYNPADPPLVLAVTSSV-----LPMTQLQDMVETRISQKI--SQVNGVGLVS--	193
Qy	995	SITMGSVRPSVEQPHKYLPMFLN-----DRPNIKCPKGGLAAYSTSV	1036
Db	194	--LAGQORPAVRV-----KLNQAASVGLDSETRVAINNANVNSAKGSLDGPTRSV	244
Qy	1037	NLTSQDGLASRFMAHYHKPLKNSQDYTEALRAARELA-----A	1074
Db	245	TL SANDQ-----MKSDDYRKLVSYKNGAPIRLADIATIEQAPENNQLGAWA	292
Qy	1075	N-----ITADLRKVPG-----TDPAEVEFPYITN-----VFYEQLTI---LPBGL	1113
Db	293	NNQALIIINQROPGVNVVETTDNIENLLPDLVSNLPKSVNVEILTDRTTIRASVKDVQ	352
Qy	1114	FMLSCLVPTFPFVSCLLGLDLSRSGILNL--LSIVMILVDVTGFWALWDISYNAVSLIN	1170
Db	353	FELGLAI-----ALUVVIVYLFKNGVATLIPSTAVPLSLVGPVAVMYFCGFSVNNITLMA	408
Qy	1171	LVSAGVMSVE----FVSHITRSPAISTKP--TWLERAKEATISMGSVAFGAVMTNLPGI	1224
Db	409	LTATGFWDDAIVVIEINISRYLERGDKPLTAALKGAGEIGFTIISLTFSLIIV--LIPL	466
Qy	1225	LVTGLAKAQLIQIFPRINMLLITLLGLLHGLVPLPVILSVGPDVNPALAEOKRAEE--	1282
Db	467	LFMGDIVGRLEFRFAITLAVAI--LISAVVSLTLPMMCARL---LKPESQIKHNRPEMAC	522
Qy	1283	-----AAVAVMVASCNHH	1295
Db	523	ERPFERKMIAYVWLKRVLNH	543
RESULT 97			
US-09-540-236-2360			
; Sequence 2360, Application US/09540236			
; Patent No. 6673910			
; GENERAL INFORMATION:			
; APPLICANT: Gary L. Breton et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXE			
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 2709.2005-001			
; CURRENT APPLICATION NUMBER: US/09/540,236			
; CURRENT FILING DATE: 2000-04-04			
; NUMBER OF SEQ ID NOS: 3840			
; SEQ ID NO 2360			
; LENGTH: 506			
; TYPE: PRT			
; ORGANISM: M.catarrhalis			
US-09-540-236-2360			
Query Match 1.6%; Score 111.5; DB 2; Length 506;			
Best Local Similarity 21.1%; Pred. No. 0.35;			
Matches 111; Conservative 70; Mismatches 167; Indels 177; Gaps 27;			
Qy	499	TLALLPANQLMGQTSQ-----VDMKDHFLYCANAPLTFKDGTL--ALSCMADYGA	548
Db	54	TL SVGLNMLLAVISQALGLVPAGVSGELFIDGFA--IFNMGVILIAALACCT----	107
Qy	549	PVPFPFLAIGCYKG-KDYSEAEALIMTFSLNN--YPAGDPRLAQAKLWEBAFLPEHRAFOR	605
Db	108	-----LAYGYFTLOPHKDELYLIMLISLIGAMLMTGAHLA-----AFFMSLELSV	155
Qy	606	RMAGMFQVFTFAERSLEDEI-----NRTTAEDL-----PIPATSYIVIF--LVISLALGS	653
Db	156	PMYGMAYTFLRSRSLSGKYLIMSATASATLLMGMALIFADAGTLLPKDLARNLLIGN	215
Qy	654	YSSWSRVWVDSKATLGLGVAVVLGAV--MAAMGFFSYIG-----IRSSLVILQV	701

Db 216 LSG-----LSVAGVVMMLAAAFKLSAAPHFHSWSDVYEGAPATAFLASVSK 264

Qy 702 VPFLVLS-----VGA-----DNIFIV-----LEYQLRRRPGCEPREVHIG 737

Db 265 VAMMALALRFFVGSMTAMPAPETILVIVVLSILMGNLLAIRQNNIKRMILAYSSIAHIG 324

Qy 738 PALGRV-----APSMILCSLSAICFFLGCALTMPMPAVRTFALTSGLAVILDFLLQMSARVA 793

Db 325 YALTALLSIGAGSLPLVSMYAV-----YALTS-----IGAFGV 358

Qy 794 LSLDSKQKQASRLDVCCKVPOELPPPGQEGELLGFFQKAYAPFLIHW-----ITRGVVL 850

Db 359 ITLMSPPRRCRTSI-----SSEADDLRFYQ-----LFWRRPVLTAULT 399

Qy 851 LLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPVYVFTTLGNVF 910

Db 400 VMLLSLAGIPL-----TAG-----FVTKF-----TIFFAAVQGMHF 430

Qy 911 SSEAGMNAICSSAGCNF-----SFTOKIQVATEPPEQSYLAIPA 950

Db 431 LA-AGMIITLGAIGLYYLRILVNLVYQPKVNLEPDAHQHGVKA 474

RESULT 98

US-09-711-164-443

Query Match 1.6%; Score 111; DB 2; Length 1025;

Best Local Similarity 22.4%; Pred. No. 1.4;

Matches 71; Conservative 50; Mismatches 148; Indels 48; Gaps 12;

GENERAL INFORMATION:

APPLICANT: Foreyth, R. Allyn

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY

FILE REFERENCE: ELITRA.008A

CURRENT APPLICATION NUMBER: US/09/711,164

PRIOR FILING DATE: 2000-11-09

PRIOR APPLICATION NUMBER: US 60/164415

PRIOR FILING DATE: 1999-11-9

NUMBER OF SEQ ID NOS: 469

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 443

LENGTH: 1025

TYPE: PRN

ORGANISM: Escherichia coli

US-09-711-164-443

Qy 996 ITMGSVRPSVEQFHKYL-----WFLNDRPNKCP-----KGGLAAYST 1034

Db 199 VSLDDVRTAVSNANVRKPKQGALEDGTHRWQIQTNDELKTAAYEYQPLIIHNNNGGAVRLGD 258

Qy 1035 SVNLTSQGVLASRFMAYHKPKNSQDYTEALRAAREL-AANITADLRKVPCTDPAP-EV 1092

Db 259 VATVDSQVDVRNAGMTNKP-----AIIIMRKUPEANIQTVDISIRAKLPELOQT 310

Qy 1093 FPYITINVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSQ-----LNLLSIVMIL 1149

Db 311 IPAAIDLQIAQDRSPITRASLEVEEQTLLISVALVILVFLFRSGRATIIIPAVSPVSL 370

Qy 1150 VDTVGFMAWDISYNAVSLINLVSAVMSVE-----FVSHITRSPAISTKPTWLERAKAT 1205

Db 371 IGTFAMYLCGFSLNLSMALTIATGTFVDDAIVVLENIAHLEAGMKP--LQAALQOT 428

Qy 1206 ISMGSAVFA--GVAMTNLPGILVGLAKAQIQLIFFFRNLNLLITLLGLLHGLVFLPVIL 1262

Db 429 REVGFTVLSMSLSLVAVFLP-LLLMGGLPGRLLREFAVTLSVAIG-ISLNVSLTLTPMMC 486

Qy 1263 SYVGPDVNPALALEQKR 1279

Db 487 GWMKASKPR---EQKR 500

RESULT 100

US-09-489-039A-7216

Query Match 1.6%; Score 111; DB 2; Length 1025;

Best Local Similarity 22.4%; Pred. No. 1.4;

Matches 71; Conservative 50; Mismatches 148; Indels 48; Gaps 12;

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7216

LENGTH: 1048

TYPE: PRN

Db 216 LSG-----LSVAGVVMMLAAAFKLSAAPHFHSWSDVYEGAPATAFLASVSK 264

Qy 702 VPFLVLS-----VGA-----DNIFIV-----LEYQLRRRPGCEPREVHIG 737

Db 265 VAMMALALRFFVGSMTAMPAPETILVIVVLSILMGNLLAIRQNNIKRMILAYSSIAHIG 324

Qy 738 PALGRV-----APSMILCSLSAICFFLGCALTMPMPAVRTFALTSGLAVILDFLLQMSARVA 793

Db 325 YALTALLSIGAGSLPLVSMYAV-----YALTS-----IGAFGV 358

Qy 794 LSLDSKQKQASRLDVCCKVPOELPPPGQEGELLGFFQKAYAPFLIHW-----ITRGVVL 850

Db 359 ITLMSPPRRCRTSI-----SSEADDLRFYQ-----LFWRRPVLTAULT 399

Qy 851 LLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPVYVFTTLGNVF 910

Db 400 VMLLSLAGIPL-----TAG-----FVTKF-----TIFFAAVQGMHF 430

Qy 911 SSEAGMNAICSSAGCNF-----SFTOKIQVATEPPEQSYLAIPA 950

Db 431 LA-AGMIITLGAIGLYYLRILVNLVYQPKVNLEPDAHQHGVKA 474

RESULT 98

US-09-711-164-443

Query Match 1.6%; Score 111; DB 2; Length 1025;

Best Local Similarity 22.4%; Pred. No. 1.4;

Matches 71; Conservative 50; Mismatches 148; Indels 48; Gaps 12;

GENERAL INFORMATION:

APPLICANT: Foreyth, R. Allyn

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY

FILE REFERENCE: ELITRA.008A

CURRENT APPLICATION NUMBER: US/09/711,164

PRIOR FILING DATE: 2000-11-09

PRIOR APPLICATION NUMBER: US 60/164415

PRIOR FILING DATE: 1999-11-9

NUMBER OF SEQ ID NOS: 469

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 443

LENGTH: 1025

TYPE: PRN

ORGANISM: Escherichia coli

US-09-711-164-443

Qy 996 ITMGSVRPSVEQFHKYL-----WFLNDRPNKCP-----KGGLAAYST 1034

Db 199 VSLDDVRTAVSNANVRKPKQGALEDGTHRWQIQTNDELKTAAYEYQPLIIHNNNGGAVRLGD 258

Qy 1035 SVNLTSQGVLASRFMAYHKPKNSQDYTEALRAAREL-AANITADLRKVPCTDPAP-EV 1092

Db 259 VATVDSQVDVRNAGMTNKP-----AIIIMRKUPEANIQTVDISIRAKLPELOQT 310

Qy 1093 FPYITINVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSQ-----LNLLSIVMIL 1149

Db 311 IPAAIDLQIAQDRSPITRASLEVEEQTLLISVALVILVFLFRSGRATIIIPAVSPVSL 370

Qy 1150 VDTVGFMAWDISYNAVSLINLVSAVMSVE-----FVSHITRSPAISTKPTWLERAKAT 1205

Db 371 IGTFAMYLCGFSLNLSMALTIATGTFVDDAIVVLENIAHLEAGMKP--LQAALQOT 428

Qy 1206 ISMGSAVFA--GVAMTNLPGILVGLAKAQIQLIFFFRNLNLLITLLGLLHGLVFLPVIL 1262

Db 429 REVGFTVLSMSLSLVAVFLP-LLLMGGLPGRLLREFAVTLSVAIG-ISLNVSLTLTPMMC 486

Qy 1263 SYVGPDVNPALALEQKR 1279

Db 487 GWMKASKPR---EQKR 500

RESULT 100

US-09-489-039A-7216

Query Match 1.6%; Score 111; DB 2; Length 1025;

Best Local Similarity 22.4%; Pred. No. 1.4;

Matches 71; Conservative 50; Mismatches 148; Indels 48; Gaps 12;

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7216

LENGTH: 1048

TYPE: PRN

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:56:52 ; Search time 189 Seconds
(without alignments)

2944.702 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MRAEGLRGWLLWALLRLAQ.....GSIKAGAGISNPLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA Main:*

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*

3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*

4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*

5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*

6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	4 US-10-621-758A-4	Sequence 4, Appli
2	6909	100.0	1332	4 US-10-663-208A-4	Sequence 4, Appli
3	6909	100.0	1332	4 US-10-646-301A-4	Sequence 4, Appli
4	6909	100.0	1332	4 US-10-736-769-4	Sequence 4, Appli
5	6896	99.8	1332	4 US-10-239-316-9	Sequence 9, Appli
6	6872.5	99.5	1359	4 US-10-621-758A-44	Sequence 44, Appl
7	6872.5	99.5	1359	4 US-10-663-208A-44	Sequence 44, Appl
8	6872.5	99.5	1359	4 US-10-646-301A-44	Sequence 44, Appl
9	6872.5	99.5	1359	4 US-10-736-769-4	Sequence 44, Appl
10	6536	94.6	1344	5 US-10-450-763-53052	Sequence 53052, A
11	5421.5	78.5	1331	4 US-10-621-758A-2	Sequence 2, Appli
12	5421.5	78.5	1331	4 US-10-663-208A-2	Sequence 2, Appli
13	5421.5	78.5	1331	4 US-10-646-301A-2	Sequence 2, Appli
14	5421.5	78.5	1331	4 US-10-736-769-2	Sequence 2, Appli
15	5407	78.3	1333	4 US-10-621-758A-12	Sequence 12, Appl
16	5407	78.3	1333	4 US-10-663-208A-12	Sequence 12, Appl
17	5407	78.3	1333	4 US-10-646-301A-12	Sequence 12, Appl
18	5407	78.3	1333	4 US-10-736-769-12	Sequence 12, Appl
19	4466	64.6	982	5 US-10-450-763-53050	Sequence 53050, A
20	2402.5	34.8	1278	4 US-10-208-731-2	Sequence 2, Appli
21	2402.5	34.8	1278	4 US-10-741-601-530	Sequence 530, App
22	2402.5	34.8	1278	5 US-10-741-600-1542	Sequence 1542, App
23	2402.5	34.8	1278	5 US-10-756-149-4924	Sequence 4924, App
24	2365	34.5	1319	4 US-10-208-731-4	Sequence 4, Appli
25	2262	32.7	1287	6 US-11-097-143-12003	Sequence 12003, A
26	1864.5	27.0	1223	6 US-11-097-143-2679	Sequence 2679, App
27	1329	19.2	1170	4 US-10-208-731-6	Sequence 6, Appli

28	1078	15.6	1296	4 US-10-208-731-9	Sequence 9, Appli
29	942	13.6	194	4 US-10-239-316-8	Sequence 8, Appli
30	900.5	13.0	492	4 US-10-424-599-211862	Sequence 211862, Sequence 189288,
31	859.5	12.4	541	4 US-10-424-599-189288	Sequence 189288,
32	651.5	9.4	1447	2 US-08-954-701A-19	Sequence 19, Appli
33	651.5	9.4	1447	3 US-09-898-533-5	Sequence 5, Appli
34	651.5	9.4	1447	3 US-09-754-032-19	Sequence 19, Appli
35	651.5	9.4	1447	4 US-10-421-446-19	Sequence 19, Appli
36	651.5	9.4	1447	4 US-10-791-844-6	Sequence 6, Appli
37	646.5	9.4	1434	2 US-08-954-701A-10	Sequence 10, Appli
38	646.5	9.4	1434	3 US-09-754-032-10	Sequence 10, Appli
39	646.5	9.4	1434	4 US-10-421-446-10	Sequence 10, Appli
40	628	9.1	1296	4 US-10-302-279-60	Sequence 60, Appli
41	598.5	8.7	933	4 US-10-415-934-3	Sequence 3, Appli
42	597	8.6	821	5 US-10-890-776A-4805	Sequence 4805, Ap
43	588	8.5	954	4 US-10-060-756A-3	Sequence 3, Appli
44	588	8.5	954	5 US-10-890-776A-3	Sequence 3, Appli
45	584	8.5	1203	3 US-09-990-046-2	Sequence 2, Appli
46	583.5	8.4	1182	3 US-09-990-046-7	Sequence 7, Appli
47	582.5	8.4	343	4 US-10-425-115-328081	Sequence 328081, Sequence 2, Appli
48	580	8.4	1203	3 US-09-909-280A-2	Sequence 2, Appli
49	579	8.4	1146	4 US-10-791-844-1	Sequence 1, Appli
50	563	8.1	836	5 US-10-890-776A-4807	Sequence 4807, Ap
51	559.5	8.1	793	5 US-10-890-776A-4806	Sequence 4806, Ap
52	539.5	7.8	846	5 US-10-890-776A-4808	Sequence 4808, Ap
53	531	7.7	1061	4 US-10-415-934-9	Sequence 9, Appli
54	531	7.7	1061	6 US-11-097-143-23451	Sequence 23451, A
55	526.5	7.6	1311	2 US-08-954-701A-4	Sequence 4, Appli
56	526.5	7.6	1311	3 US-09-754-032-4	Sequence 4, Appli
57	526.5	7.6	1311	4 US-10-421-446-4	Sequence 4, Appli
58	514	7.4	1405	4 US-10-369-493-5464	Sequence 5464, Ap
59	474	6.9	1286	3 US-09-898-533-3	Sequence 3, Appli
60	474	6.9	1286	6 US-11-097-143-4068	Sequence 4068, Ap
61	464.5	6.7	1299	2 US-08-954-771-48	Sequence 48, Appli
62	464.5	6.7	1299	5 US-10-647-654-48	Sequence 48, Appli
63	464.5	6.7	1299	5 US-10-835-517-48	Sequence 48, Appli
64	442.5	6.4	906	5 US-10-890-776A-4809	Sequence 4809, Ap
65	438.5	6.3	1285	2 US-08-954-701A-6	Sequence 6, Appli
66	438.5	6.3	1285	3 US-09-754-032-6	Sequence 6, Appli
67	438.5	6.3	1285	4 US-10-421-446-6	Sequence 6, Appli
68	379	5.1	714	4 US-10-369-493-5410	Sequence 5410, Ap
69	351.5	5.1	767	4 US-10-060-756A-6	Sequence 6, Appli
70	351.5	5.1	767	5 US-10-890-776A-6	Sequence 6, Appli
71	339	4.9	783	4 US-10-094-749-1772	Sequence 1772, Ap
72	310.5	4.5	648	4 US-10-060-756A-10	Sequence 10, Appli
73	310.5	4.5	648	5 US-10-890-776A-10	Sequence 10, Appli
74	306.5	4.4	648	4 US-10-060-756A-4799	Sequence 4799, Ap
75	306.5	4.4	648	5 US-10-890-776A-4799	Sequence 4799, Ap
76	291	4.2	61	3 US-09-864-761-34826	Sequence 34826, A
77	283.5	4.1	136	4 US-10-425-115-277863	Sequence 277863, Sequence 277864,
78	272.5	3.9	132	4 US-10-425-115-277864	Sequence 8, Appli
79	266.5	3.9	792	4 US-10-415-934-8	Sequence 8, Appli
80	264.5	3.8	229	4 US-10-424-599-181755	Sequence 181755, Sequence 201210,
81	259	3.7	267	4 US-10-425-115-201210	Sequence 1316, Ap
82	252	3.6	505	4 US-10-408-765A-1316	Sequence 1316, Ap
83	245.5	3.6	1491	4 US-10-432-613-2	Sequence 2, Appli
84	230	3.3	785	4 US-10-415-934-2	Sequence 2, Appli
85	225	3.3	542	4 US-10-017-161-2024	Sequence 2024, Ap
86	225	3.3	542	4 US-10-292-798-1670	Sequence 1670, Ap
87	225	3.3	563	4 US-10-415-934-14	Sequence 14, Appli
88	214.5	3.1	1218	4 US-10-432-613-3	Sequence 3, Appli
89	214.5	3.1	1218	6 US-11-097-143-3123	Sequence 3123, Ap
90	212.5	3.1	1218	4 US-10-168-428-2	Sequence 2, Appli
91	210.5	3.0	1203	4 US-10-094-749-2651	Sequence 2651, Ap
92	202	2.9	1124	4 US-10-311-623-3	Sequence 3, Appli
93	197.5	2.9	1561	4 US-10-085-198-168	Sequence 168, App
94	195.5	2.8	1087	4 US-10-239-316-17	Sequence 17, Appli
95	195.5	2.8	1392	4 US-10-239-316-40	Sequence 40, Appli
96	190	2.8	100	4 US-10-437-963-102679	Sequence 102679, Sequence 3, Appli
97	187.5	2.7	931	4 US-10-168-428-3	Sequence 3, Appli
98	177	2.6	121	4 US-10-425-115-222635	Sequence 222635, Sequence 267, App
99	174	2.5	1176	4 US-10-041-018-267	Sequence 14, Appli
100	170	2.5	101	4 US-10-060-756A-14	

ALIGNMENTS

RESULT 1

US-10-621-758A-4

; Sequence 4, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-621-758A-4

Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGRGWLLWALLRLAQSEPTTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MAEAGRGWLLWALLRLAQSEPTTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLTLOKICPRLTYGNTQACCSAKQLVLSLEASLSITKALLTRCPACSNF 120
DB 61 TPARKITGDHLLTLOKICPRLTYGNTQACCSAKQLVLSLEASLSITKALLTRCPACSNF 120
QY 121 VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
DB 121 VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTMCVYGSALCNAQRWLNFGDTGNGLAFLDITFHLLPEQAVGSGIQPLNEGV 240
DB 181 ATLAVGTMCVYGSALCNAQRWLNFGDTGNGLAFLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNESQGDVATCSQDCDCAAPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESQGDVATCSQDCDCAAPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSNVDPKGTSLSDKLSPSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
DB 301 LLVGFRVAPARDKSNVDPKGTSLSDKLSPSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDQHPFPRTNQVILITAPNRSY 420
DB 361 IPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDQHPFPRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPNKPSGILDLLELLELQERLRLHQLVWSPEAQRNLSQDICYAPLNPDNT 480
DB 421 RYDLSLLGPNKPSGILDLLELLELQERLRLHQLVWSPEAQRNLSQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCHADYGAPVFPFLAIGGKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
DB 541 SCHADYGAPVFPFLAIGGKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSSWSRV 660

RESULT 2

US-10-663-208A-4

; Sequence 4, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1

DB 601 RAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSSWSRV 660
QY 661 MVDKATILGLGGVAVVILGAVNAANGPFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MVDKATILGLGGVAVVILGAVNAANGPFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFLGLATPMPAVRTFALTSGLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFLGLATPMPAVRTFALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVXPQELPPPPQGGSGLLGFPQKAYAPPL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVXPQELPPPPQGGSGLLGFPQKAYAPPL 840
QY 841 LHWITRGVVLVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
DB 841 LHWITRGVVLVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVITTLGYNFSSEAGMNAICSSAGCANNFSTQKIQYATEFFPEQSYLEIAPASSWVDDFIDW 960
DB 901 YFVITTLGYNFSSEAGMNAICSSAGCANNFSTQKIQYATEFFPEQSYLEIAPASSWVDDFIDW 960
QY 961 LTPSSCCRLXIISGPNKDKFCPSTVNSLNCNCHMSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
DB 961 LTPSSCCRLXIISGPNKDKFCPSTVNSLNCNCHMSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVPPYTIINVFYEQVLTILPGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAPFVPPYTIINVFYEQVLTILPGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVGPMAWDISYNVAVSLINLVSAVMSVPEVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSIVMLVDTVGPMAWDISYNVAVSLINLVSAVMSVPEVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISGSAVFAVAVANTNLPGLIIVLGLAKAQLIQIIFFRNLNLTITLGLHLGLVFLPV 1260
DB 1201 AKEATISGSAVFAVAVANTNLPGLIIVLGLAKAQLIQIIFFRNLNLTITLGLHLGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSRVSTADNIYVNHSPGSIKAGAGA 1320
DB 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSRVSTADNIYVNHSPGSIKAGAGA 1320
QY 1321 ISNPLPNNNGRQF 1332
DB 1321 ISNPLPNNNGRQF 1332


```
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-208A-4

Query Match      100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
QY 181 ATLVGTMCGVYGSALCNAQRLNFGDGTGNGLAPLDITFHLEPGQAVGSGIQLNEG 240
DB 181 ATLVGTMCGVYGSALCNAQRLNFGDGTGNGLAPLDITFHLEPGQAVGSGIQLNEG 240
QY 241 ARCNEISQDDVATCSCQDCAAPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNEISQDDVATCSCQDCAAPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFPGMGWTVASWPLTLVLVSV 360
DB 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFPGMGWTVASWPLTLVLVSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDOHGFPPFTNQVILTAENRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDOHGFPPFTNQVILTAENRSY 420
QY 421 RYDLSLLGPKNPSGILDLLELLELQERLHLQVWSPQARNSLQDICVAPLNPNPT 480
DB 421 RYDLSLLGPKNPSGILDLLELLELQERLHLQVWSPQARNSLQDICVAPLNPNPT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTPKQGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTPKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGYKDYSEALIMTFSLNYPAGDPRLAQAKLWEFAFLBEM 600
DB 541 SCWADYGAPVFPFLAIGYKGYKDYSEALIMTFSLNYPAGDPRLAQAKLWEFAFLBEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWSRV 660
DB 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWSRV 660
QY 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 780
DB 721 YORLPRRPGEPREVIHGRALGVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPQCGEGLLGPPQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPQCGEGLLGPPQKAYAPFL 840
QY 841 LHWITRGVVLFLFLFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLFLFLFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGNFSSEAGMNAICSSAGCNFSTQKIQYATEFPFQSYLAIPASSWVDDFDIW 960
DB 901 YFVTTILGNFSSEAGMNAICSSAGCNFSTQKIQYATEFPFQSYLAIPASSWVDDFDIW 960

; Sequence 4, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-4

Query Match      100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLITKALLTRCPACSNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
QY 181 ATLVGTMCGVYGSALCNAQRLNFGDGTGNGLAPLDITFHLEPGQAVGSGIQLNEG 240
DB 181 ATLVGTMCGVYGSALCNAQRLNFGDGTGNGLAPLDITFHLEPGQAVGSGIQLNEG 240
QY 241 ARCNEISQDDVATCSCQDCAAPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNEISQDDVATCSCQDCAAPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFPGMGWTVASWPLTLVLVSV 360
DB 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFPGMGWTVASWPLTLVLVSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDOHGFPPFTNQVILTAENRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDOHGFPPFTNQVILTAENRSY 420
QY 421 RYDLSLLGPKNPSGILDLLELLELQERLHLQVWSPQARNSLQDICVAPLNPNPT 480
DB 421 RYDLSLLGPKNPSGILDLLELLELQERLHLQVWSPQARNSLQDICVAPLNPNPT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTPKQGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTPKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGYKDYSEALIMTFSLNYPAGDPRLAQAKLWEFAFLBEM 600
DB 541 SCWADYGAPVFPFLAIGYKGYKDYSEALIMTFSLNYPAGDPRLAQAKLWEFAFLBEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWSRV 660
DB 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWSRV 660
QY 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 780
DB 721 YORLPRRPGEPREVIHGRALGVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPQCGEGLLGPPQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKPQELPPQCGEGLLGPPQKAYAPFL 840
QY 841 LHWITRGVVLFLFLFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLFLFLFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGNFSSEAGMNAICSSAGCNFSTQKIQYATEFPFQSYLAIPASSWVDDFDIW 960
DB 901 YFVTTILGNFSSEAGMNAICSSAGCNFSTQKIQYATEFPFQSYLAIPASSWVDDFDIW 960
```

Db 181 ATLAGVTGCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFRTNQVILTAPNRSSY 420
Qy 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLQWSPQARNISLODICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLQWSPQARNISLODICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
Qy 541 SCMADYGAPVFFFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
Db 541 SCMADYGAPVFFFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
Qy 601 RAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWRV 660
Db 601 RAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWRV 660
Qy 661 MYDSKATILGLGVAVVILGAVMAAMGFESYLGIRSSILVILQVVPFLVLSVGADNIFIVLE 720
Db 661 MYDSKATILGLGVAVVILGAVMAAMGFESYLGIRSSILVILQVVPFLVLSVGADNIFIVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCVKQOEI LPPQGGGLLIGFPQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCVKQOEI LPPQGGGLLIGFPQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA PV 900
Db 841 LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA PV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNFFSFTQKI QVATEFPFQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFFSFTQKI QVATEFPFQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Qy 1021 NIKCPKGGAAVYSTSVNLSDGVLASRFMAVHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGGAAVYSTSVNLSDGVLASRFMAVHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPDGPDAFEVFPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDRSGLL 1140
Db 1081 RKVPDGPDAFEVFPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDRSGLL 1140
Qy 1141 NLLSIVMILVDVTFGMALWDISVNAVSLINLSVAGMSEVESHITRSFAISTKPTWLER 1200
Db 1141 NLLSIVMILVDVTFGMALWDISVNAVSLINLSVAGMSEVESHITRSFAISTKPTWLER 1200
Qy 1201 AKEATISMSGSAVPAVAMTNLPGILVLGLAKAQI QIFFRNLNLTLLGLLHGLVFLPV 1260
Db 1201 AKEATISMSGSAVPAVAMTNLPGILVLGLAKAQI QIFFRNLNLTLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPDVNPALALQKRAEEAAVAVMVASCPNHPFSRVSTADNIYVNHSPFSGSIKAGA 1320
Db 1261 ILSYVGPDVNPALALQKRAEEAAVAVMVASCPNHPFSRVSTADNIYVNHSPFSGSIKAGA 1320

Qy 1321 ISNPLPNNGRQF 1332
Db 1321 ISNPLPNNGRQF 1332

RESULT 4

US-10-736-769-4
; Sequence 4, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J801603-K3-US
; CURRENT APPLICATION NUMBER: US/10736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-4

Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLLLRLAQSEPTTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLLLRLAQSEPTTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Qy 61 TPARKITGDHILLQKICPRLTYGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILLQKICPRLTYGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAGVTGCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAGVTGCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFRTNQVILTAPNRSSY 420
Qy 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLQWSPQARNISLODICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLQWSPQARNISLODICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540

```
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTTKDGDTALAL 540
Qy 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Qy 601 RAQORMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV 660
Db 601 RAQORMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV 660
Qy 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Qy 721 YQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKVQDELPPGQEGLLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKVQDELPPGQEGLLLGFFQKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTLYGNFSSEAGMNAICSSAGCNCFSTQKIQVATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTLYGNFSSEAGMNAICSSAGCNCFSTQKIQVATEFPQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCIKNMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCIKNMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGGAAVSTSVNLTSQGVLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGGAAVSTSVNLTSQGVLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVFPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIVMILVDTGVFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTGVFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
Qy 1201 AKEATISMGSAFVAGVAMTNLPGILVLGAKAQLIQIIFPRNLNLTLLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAFVAGVAMTNLPGILVLGAKAQLIQIIFPRNLNLTLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPSGSIKGAGA 1320
Db 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPSGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 5
US-10-239-316-9
; Sequence 9, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1e1 Protein, Process for Producing The Same And
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
```

```
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
; US-10-239-316-9

Query Match          99.8%; Score 6896; DB 4; Length 1332;
Best Local Similarity 99.8%; Pred No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTENGLAPLDITFHLLEPCQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTENGLAPLDITFHLLEPCQAVGSGIQPLNEGV 240
Qy 241 ARCNSQGDVATCQDCCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGDVATCQDCCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTTLGQFFQGMGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTTLGQFFQGMGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVTELTTPDELWSAPNSQARSEKAFHQHGFPPFRTNQTIVITAPNRSY 420
Db 361 IPVVALAAGLVTELTTPDELWSAPNSQARSEKAFHQHGFPPFRTNQTIVITAPNRSY 420
Qy 421 RYDLSLLGPKNPSGILDLDLLELEQLERHLQVMSPEAQRTSLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDLLELEQLERHLQVMSPEAQRTSLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTTKDGDTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTTKDGDTALAL 540
Qy 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Qy 601 RAQORMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV 660
Db 601 RAQORMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV 660
Qy 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Qy 721 YQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKVQDELPPGQEGLLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKVQDELPPGQEGLLLGFFQKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
```

```
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIOYATEPPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIOYATEPPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCSTVNSLCLNCKMSITMGSRVPSVEQPHKYLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCSTVNSLCLNCKMSITMGSRVPSVEQPHKYLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLAIRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLAIRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDLRSGLL 1140
QY 1141 NLLSVMTILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSVMTILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAVAMTNLPGLILVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVFAVAMTNLPGLILVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVNPALALEOKAEAEAAVAVMVASCPNHPRSVSTADNIYVNHSPSGSKGAGA 1320
DB 1261 ILSYVGPVNPALALEOKAEAEAAVAVMVASCPNHPRSVSTADNIYVNHSPSGSKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 6
US-10-621-758A-44
; Sequence 44, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Juquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-621-758A-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLWGMLLWALLRLAQSEPYTTTHORGYCAFYDECCKNPGLSGSLMTLSNVSCLSN 60
DB 1 MAEAGLWGMLLWALLRLAQSEPYTTTHORGYCAFYDECCKNPGLSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPLRYTGPNTOACCSAKQLVLEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPLRYTGPNTOACCSAKQLVLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSFNQSLFINVTRVLAQAGQLPAVAVAEAFYQHSFASQSDYSCSRVRPAA 180
DB 121 VNLHCHTCSFNQSLFINVTRVLAQAGQLPAVAVAEAFYQHSFASQSDYSCSRVRPAA 180
QY 181 ATLAVGTMCGVYGSAICNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
```

```
DB 181 ATLAVGTMCGVYGSAICNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
QY 241 ARCNEQODDVATCSCODCAASCPAIARPOALDSTFVLGOMPGSLVLIILICSVFVAVTTI 300
DB 241 ARCNEQODDVATCSCODCAASCPAIARPOALDSTFVLGOMPGSLVLIILICSVFVAVTTI 300
QY 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQHPGPPFRFNQVTLITAPNRSY 420
DB 361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQHPGPPFRFNQVTLITAPNRSY 420
QY 421 RYDSLILGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
DB 421 RYDSLILGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKQHFLYCANAPLTFKDGTAALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKQHFLYCANAPLTFKDGTAALAL 540
QY 541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAEFLSEEM 600
DB 541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAEFLSEEM 600
QY 601 RAFQRRMAGMFQVTFMAERSLEDEINRTTAEADLPFATSYIVIFLYISLALGYSWSNRV 660
DB 601 RAFQRRMAGMFQVTFMAERSLEDEINRTTAEADLPFATSYIVIFLYISLALGYSWSNRV 660
QY 661 MVDKATLGLGGAVALGAVMAAMGFFSYLGISSVLILQVVPFLVLSVSGADNIFIVLE 720
DB 661 MVDKATLGLGGAVALGAVMAAMGFFSYLGISSVLILQVVPFLVLSVSGADNIFIVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMILLCSLSEAI CFFLICALTPMPAVRIFALTSLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMILLCSLSEAI CFFLICALTPMPAVRIFALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVKPOELPPGQGGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVKPOELPPGQGGGLLGFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
DB 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIOYATEPPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIOYATEPPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCSTVNSLCLNCKMSITMGSRVPSVEQPHKYLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCSTVNSLCLNCKMSITMGSRVPSVEQPHKYLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLAIRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLAIRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 KPLKNSQDYTEALRAARELANITADLRAKVPCTDPAFEVFPYTTINVFYEQYLTILPEGL 1140
DB 1081 KPLKNSQDYTEALRAARELANITADLRAKVPCTDPAFEVFPYTTINVFYEQYLTILPEGL 1140
QY 1141 FMLSCLVPTPAVSCLLGLDLRSGLLNLISVILVDTVGFMAWLDISYNAVSLINLVS 1200
DB 1141 FMLSCLVPTPAVSCLLGLDLRSGLLNLISVILVDTVGFMAWLDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAVAMTNLPGLILVGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAVAMTNLPGLILVGLAKAQ 1260
QY 1234 LQIIFPRNLNLLITLGLLHGLVFLPVILSYVGPVNPALALEOKAEAEAAVAVMVASCP 1293
DB 1234 LQIIFPRNLNLLITLGLLHGLVFLPVILSYVGPVNPALALEOKAEAEAAVAVMVASCP 1293
```

Db 1261 LIQIPFRLNLLITLLGLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVASCP 1320
QY 1294 NHPSRVSTADNIYNHSPFGSIKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYNHSPFGSIKGAGAI SNFLPNNGRQF 1359

RESULT 7

US-10-663-208A-44
; Sequence 44, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; PRIOR FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-208A-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLQKICPLRYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPLRYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFAEQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFAEQSDSCSRVRPAA 180
QY 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNSQGDVATCSQDCAASCPAIAAPQALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGDVATCSQDCAASCPAIAAPQALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSNVDPKGTSLSDKLSFSFTHLLGQPFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSNVDPKGTSLSDKLSFSFTHLLGQPFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTDDVELWSAPNSQARSEKAFPHQFPGFFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVTELTDDVELWSAPNSQARSEKAFPHQFPGFFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPNKPSGILDLDLLELLELQERHLQVWSPQARNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPNKPSGILDLDLLELLELQERHLQVWSPQARNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTFKDTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTFKDTALAL 540

QY 541 SCMDYGAVPFPFLAIGGYKGKDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAELEBM 600
Db 541 SCMDYGAVPFPFLAIGGYKGKDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAELEBM 600
QY 601 RAFQRMAGMFQVTTFAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALAGSYSSWSRV 660
Db 601 RAFQRMAGMFQVTTFAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALAGSYSSWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVSLGADNIRIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVSLGADNIRIFVLE 720
QY 721 YQRLPRRFPCEPREVHIGRALGRVAPSMILCSLSEAI CFFELGALTMPA VRTTALTSLGLAV 780
Db 721 YQRLPRRFPCEPREVHIGRALGRVAPSMILCSLSEAI CFFELGALTMPA VRTTALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCCKVQBELPPPGQGGEGLLIGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCCKVQBELPPPGQGGEGLLIGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFELFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFELFLNRYFEVGAPV 900
QY 901 YFVTTLYGNFSSEAGMNAICSSAGCWNFSFTQKIQYATEFPPEQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTLYGNFSSEAGMNAICSSAGCWNFSFTQKIQYATEFPPEQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNLSNCLKNCMSITMGSVRPSVEQPHKYLPWFNLDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSVTNLSNCLKNCMSITMGSVRPSVEQPHKYLPWFNLDRP 1020
QY 1021 NIKPKGGLAAVSTSVNLTSDQVLI-----ASRFMAVH 1053
Db 1021 NIKPKGGLAAVSTSVNLTSDQVLI-----ASRFMAVH 1053
QY 1054 KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYTTITNVEQYLTILPEGL 1113
Db 1054 KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYTTITNVEQYLTILPEGL 1113
QY 1114 PMLSCLVPTFAVSCLLIGLRLSGLNLLSITVILVDTVGPALWDSINAVSLINLVS 1173
Db 1141 PMLSCLVPTFAVSCLLIGLRLSGLNLLSITVILVDTVGPALWDSINAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIPFRLNLLITLLGLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVASCP 1293
Db 1261 LIQIPFRLNLLITLLGLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVASCP 1320
QY 1294 NHPSRVSTADNIYNHSPFGSIKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYNHSPFGSIKGAGAI SNFLPNNGRQF 1359

RESULT 8

US-10-646-301A-44
; Sequence 44, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19

		; PRIOR APPLICATION NUMBER: 10/621,758			
		; PRIOR FILING DATE: 2003-07-17			
		; NUMBER OF SEQ ID NOS: 50			
		; SOFTWARE: PatentIn version 3.1			
		; SEQ ID NO 44			
		; LENGTH: 1359			
		; TYPE: PRT			
		; ORGANISM: Homo sapiens			
		US-10-646-301A-44			
		Query Match 99.5%; Score 6872.5; DB 4; Length 1359;			
		Best Local Similarity 97.9%; Pred. No. 0;			
		Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;			
Qy	1	MAEAGLRGWLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN	60		
Db	1	MAEAGLRGWLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN	60		
Qy	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Qy	121	VNLHCHNTCSPNQSIFINTRVAQLGAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA	180		
Db	121	VNLHCHNTCSPNQSIFINTRVAQLGAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA	180		
Qy	181	ATLAVGTWCGVYGSAALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240		
Db	181	ATLAVGTWCGVYGSAALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240		
Qy	241	ARCNSQGGDVATCQDCAACSPAIARPOALDSTFYLGOMPGSLVLIILCSVFVVTI	300		
Db	241	ARCNSQGGDVATCQDCAACSPAIARPOALDSTFYLGOMPGSLVLIILCSVFVVTI	300		
Qy	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTLVLVS	360		
Db	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTLVLVS	360		
Qy	361	IPVVALAAGLVFTELTTDPVELWSPNSQARSEKAFHDQHGFPFRTNQVILTAENRSY	420		
Db	361	IPVVALAAGLVFTELTTDPVELWSPNSQARSEKAFHDQHGFPFRTNQVILTAENRSY	420		
Qy	421	RYDSLGLGPNFSGILDLLELLELQERLRLQVSPQARNISLODICVAPLNPDNT	480		
Db	421	RYDSLGLGPNFSGILDLLELLELQERLRLQVSPQARNISLODICVAPLNPDNT	480		
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDVDKDHFLYCANAPLTFKDGITALAL	540		
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDVDKDHFLYCANAPLTFKDGITALAL	540		
Qy	541	SCMADYGAPVPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEEAPLEEM	600		
Db	541	SCMADYGAPVPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEEAPLEEM	600		
Qy	601	RAFQRMAGNFQVTFMAERSLEDEINRTAEDLPIEATSVIVFLYISLALGYSWSRV	660		
Db	601	RAFQRMAGNFQVTFMAERSLEDEINRTAEDLPIEATSVIVFLYISLALGYSWSRV	660		
Qy	661	MVDSKATLGGVAVVLGAVMAAGPFYSYIGIRSSILVILQVVPFLVLSGADNIFIFVLE	720		
Db	661	MVDSKATLGGVAVVLGAVMAAGPFYSYIGIRSSILVILQVVPFLVLSGADNIFIFVLE	720		
Qy	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780		
Db	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780		
Qy	781	ILDFLLQMSAFVALLSDSKQASRLDVCCVKPQELPPPGQEGLLGFFOKAYAPFL	840		
Db	781	ILDFLLQMSAFVALLSDSKQASRLDVCCVKPQELPPPGQEGLLGFFOKAYAPFL	840		
Qy	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900		
Db	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900		
		Query Match 99.5%; Score 6872.5; DB 4; Length 1359;			
		Best Local Similarity 97.9%; Pred. No. 0;			
		Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;			
Qy	1	MAEAGLRGWLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN	60		
Db	1	MAEAGLRGWLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN	60		
Qy	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Qy	121	VNLHCHNTCSPNQSIFINTRVAQLGAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA	180		
Db	121	VNLHCHNTCSPNQSIFINTRVAQLGAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA	180		
Qy	181	ATLAVGTWCGVYGSAALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240		
Db	181	ATLAVGTWCGVYGSAALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240		
Qy	241	ARCNSQGGDVATCQDCAACSPAIARPOALDSTFYLGOMPGSLVLIILCSVFVVTI	300		
Db	241	ARCNSQGGDVATCQDCAACSPAIARPOALDSTFYLGOMPGSLVLIILCSVFVVTI	300		
Qy	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTLVLVS	360		
Db	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTLVLVS	360		
Qy	361	IPVVALAAGLVFTELTTDPVELWSPNSQARSEKAFHDQHGFPFRTNQVILTAENRSY	420		
Db	361	IPVVALAAGLVFTELTTDPVELWSPNSQARSEKAFHDQHGFPFRTNQVILTAENRSY	420		
Qy	421	RYDSLGLGPNFSGILDLLELLELQERLRLQVSPQARNISLODICVAPLNPDNT	480		
Db	421	RYDSLGLGPNFSGILDLLELLELQERLRLQVSPQARNISLODICVAPLNPDNT	480		
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDVDKDHFLYCANAPLTFKDGITALAL	540		
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDVDKDHFLYCANAPLTFKDGITALAL	540		
Qy	541	SCMADYGAPVPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEEAPLEEM	600		
Db	541	SCMADYGAPVPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEEAPLEEM	600		
Qy	601	RAFQRMAGNFQVTFMAERSLEDEINRTAEDLPIEATSVIVFLYISLALGYSWSRV	660		
Db	601	RAFQRMAGNFQVTFMAERSLEDEINRTAEDLPIEATSVIVFLYISLALGYSWSRV	660		
Qy	661	MVDSKATLGGVAVVLGAVMAAGPFYSYIGIRSSILVILQVVPFLVLSGADNIFIFVLE	720		
Db	661	MVDSKATLGGVAVVLGAVMAAGPFYSYIGIRSSILVILQVVPFLVLSGADNIFIFVLE	720		
Qy	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780		
Db	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780		
Qy	781	ILDFLLQMSAFVALLSDSKQASRLDVCCVKPQELPPPGQEGLLGFFOKAYAPFL	840		
Db	781	ILDFLLQMSAFVALLSDSKQASRLDVCCVKPQELPPPGQEGLLGFFOKAYAPFL	840		
Qy	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900		
Db	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900		

RESULT 9
US-10-736-769-44
; Sequence 44, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION: Scott W
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-44

Db 61 TPARKITGDHLLILLQKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
QY 181 ATLAVTGTCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
Db 181 ATLAVTGTCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
QY 241 ARCNEQSGDDVATCSQDCQCAASCPAIPARPOALDSTFYLGMQPSGLVLIILCSVFAVVTI 300
Db 241 ARCNEQSGDDVATCSQDCQCAASCPAIPARPOALDSTFYLGMQPSGLVLIILCSVFAVVTI 300
QY 301 LLVGRFVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSV 360
Db 301 LLVGRFVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDQHPGPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDQHPGPFRTNQVILITAPNRSY 420
QY 421 RYDSSLILGPKNFGIILDLLELLELQERLRLHQLWSPEAORNLSDICVAPLNPDNT 480
Db 421 RYDSSLILGPKNFGIILDLLELLELQERLRLHQLWSPEAORNLSDICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEARALIMTFSLNNYPAGDPRLAQAKLWEAPLEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEARALIMTFSLNNYPAGDPRLAQAKLWEAPLEM 600
QY 601 RAPORMAGMFWQVTFABSLDEINRTTAEDLPATSYIYIYIYISIALGYSWSRV 660
Db 601 RAPORMAGMFWQVTFABSLDEINRTTAEDLPATSYIYIYIYISIALGYSWSRV 660
QY 661 MYDSKATLGLGVAVVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVGADNIPFLVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVGADNIPFLVLE 720
QY 721 YQRLPRRPGEPREVIHIGALGRVAFSMLLCSLSEACFPFLGALTMPAVRTFALTSGLA 780
Db 721 YQRLPRRPGEPREVIHIGALGRVAFSMLLCSLSEACFPFLGALTMPAVRTFALTSGLA 780
QY 781 ILDFELLQMSAFVALLSLDSKQREASRLDYCCCKPQELPPQCGGLLGFQKAYAPPL 840
Db 781 ILDFELLQMSAFVALLSLDSKQREASRLDYCCCKPQELPPQCGGLLGFQKAYAPPL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLPLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLPLNRYFEVGAPV 900
QY 901 YFVTTLYGNFSSBAGMNAICSSAGCNGNFSFTQKIYATFEPFQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTLYGNFSSBAGMNAICSSAGCNGNFSFTQKIYATFEPFQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSRPVSVEQFHKYLPWFILNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSRPVSVEQFHKYLPWFILNDRP 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDQVL-----ASRFMAVH 1053
Db 1021 NIKCPKGGLAAYSTSVNLTSDQVL-----ASRFMAVH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTINVFYEQILTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTINVFYEQILTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDRLSGLLNLLSIVMILVDTGPMALWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDRLSGLLNLLSIVMILVDTGPMALWDISYNAVSLINLVS 1200

RESULT 10

US-10-450-763-53052
; Sequence 53052, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53052
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (901)..(917)
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX.
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1344)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53052

Query Match 94.6%; Score 6536; DB 5; Length 1344;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLQKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLQKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
QY 181 ATLAVTGTCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
Db 181 ATLAVTGTCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
QY 241 ARCNEQSGDDVATCSQDCQCAASCP--AATARPQALDSTFYLGMQPSGLVLIILCSVFAV 298
Db 241 ARCNEQSGDDVATCSQDCQCAASCPXPAPRSTPTSTW--RQMPGSLVLIILCSVFAV 298
QY 299 TLLVGRFVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILV 358
Db 299 TLLVGRFVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILV 358

Db 299 TILLVGRVAPARDKSKMVDPKGTSLSDKLSFSHTTLTGQFFQGWGTTWASWPLTILVL 358
Qy 359 SVIPVVALAAGLVFTLTTDPVELWSPNSQARSEKAFHDHGFPPFRFTNOVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTLTTDPVELWSPNSQARSEKAFHDHGFPPFRFTNOVILTAPNRS 418
Qy 419 SYRVDLSLLGPKNPSGIILDLLELLELRLHQLWSPQAORNIISLQDICYAPLNPD 478
Db 419 SYRVDLSLLGPKNPSGIILDLLELLELRLHQLWSPQAORNIISLQDICYAPLNPD 478
Qy 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTA 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTA 538
Qy 539 ALSMADYGAVPFPPLAIGYKGDYSEABALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 598
Db 539 ALSMADYGAVPFPPLAIGYKGDYSEABALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 598
Qy 599 EMRAFQRMAGMFOVTTAERSLEDEINRTTAEDLPIFATSVIVIFLYISALGSYSWS 658
Db 599 EMRAFQRMAGMFOVTTAERSLEDEINRTTAEDLPIFATSVIVIFLYISALGSYSWS 658
Qy 659 RVWVDSKATLGLGGVAVVLGAWMAAGFFSYLGIRSSVLIVQVFPFLVLSVGADNIFIV 718
Db 659 RVWVDSKATLGLGGVAVVLGAWMAAGFFSYLGIRSSVLIVQVFPFLVLSVGADNIFIV 718
Qy 719 LEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSG 778
Db 719 LEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSG 778
Qy 779 AVILDFLLQMSAFVALLSDSKROASRLDVCCKVQBELPPGQEGELLGFFOKAYAP 838
Db 779 AVILDFLLQMSAFVALLSDSKROASRLDVCCKVQBELPPGQEGELLGFFOKAYAP 838
Qy 839 FLLHWITRGW----LFLFLFGLVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYF 894
Db 839 FLLHWITRGWVVPVPSQLLLFLFGLVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYF 898
Qy 895 EVGAPVYFVTTILGYNPSSEAGMNAICSSAGCNFFSTQKIQVATEPPEQSYLAIPASSWV 954
Db 895 EVGAPVYFVTTILGYNPSSEAGMNAICSSAGCNFFSTQKIQVATEPPEQSYLAIPASSWV 958
Qy 955 DFDIDLWLTSSCCRLYISGPNKDKCPSVTNSLNCNKMSITMGSVRPSVSOHQKYLWP 1014
Db 959 DFDIDLWLTSSCCRLYISGPNKDKCPSVTNSLNCNKMSITMGSVRPSVSOHQKYLWP 1018
Qy 1015 FLNDRNIIKCPKGLAAYSTSVNLTSQVLAASRFMAHKLKNSQDYTEALRAARELAA 1074
Db 1019 FLNDRNIIKCPKGLAAYSTSVNLTSQVLAASRFMAHKLKNSQDYTEALRAARELAA 1078
Qy 1075 NITADLRKVPGTDPAEVFPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLD 1134
Db 1079 NITADLRKVPGTDPAEVFPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLD 1138
Qy 1135 LRSGLLNLSIVMILVDYVGFMAWLDI SYNNAVSLINLVSAGMSVEFVSHITRSPAISTK 1194
Db 1139 LRSGLLNLSIVMILVDYVGFMAWLDI SYNNAVSLINLVSAGMSVEFVSHITRSPAISTK 1198
Qy 1195 PTLWLERAKATISMGSAVAGVANTLPGILVLGLAKAQLIQIFPFLMLLTITLGLLHG 1254
Db 1199 PTLWLERAKATISMGSAVAGVANTLPGILVLGLAKAQLIQIFPFLMLLTITLGLLHG 1258
Qy 1255 LVFLPVLISYVGPDPVNPALAEQKRAEEA-----VAAMVVASCPNHPSPRVST 1301
Db 1259 LVFLPVLISYVGPDPVNPALAEQKRAEEGGSGHGGLAQITPPESPQLTTSMT 1313

RESULT 11

US-10-621-758A-2

; Sequence 2, Application US/10621758A

; Publication No. US20040093629A1

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; PRIORITY FILING DATE: 2003-07-17
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2

Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECCKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECCKNPELSGSLMTLSNVCSLSN 59
Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTQ-ACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILLOKICPRLYTGPNTQ-ACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
Qy 120 FVNHLCHNTCPNOSLFINVTRVAQAGQLPVAVVAYEAFYQHSFAESYSCSRVRVPA 179
Db 120 FVNHLCHNTCPNOSLFINVTRVAQAGQLPVAVVAYEAFYQHSFAESYSCSRVRVPA 179
Qy 180 AATLAVGTCMGVYSALCNQARWLNFGDGTGNGLAPLDITLHLLPQAVGSGIOLNEG 239
Db 180 AATLAVGTCMGVYSALCNQARWLNFGDGTGNGLAPLDITLHLLPQAVGSGIOLNEG 239
Qy 240 VARNCSOGDDVATCSCODCAASCPAIARPOALDSTFYLQWPGSLVLIILICSVFVVT 299
Db 240 VARNCSOGDDVATCSCODCAASCPAIARPOALDSTFYLQWPGSLVLIILICSVFVVT 299
Qy 300 ILLVGRVAPARDKSKMVDPKGTSLSDKLSFSHTTLTGQFFQGWGTTWASWPLTILVL 359
Db 300 ILLVGRVAPARDKSKMVDPKGTSLSDKLSFSHTTLTGQFFQGWGTTWASWPLTILVL 359
Qy 360 VIPVVALAAGLVFTLTTDPVELWSPNSQARSEKAFHDHGFPPFRFTNOVILTAPNRS 419
Db 360 VIPVVALAAGLVFTLTTDPVELWSPNSQARSEKAFHDHGFPPFRFTNOVILTAPNRS 419
Qy 420 YRYSLLLLGPKNPSGIILDLLELLELRLHQLWSPQAORNIISLQDICYAPLNPD 479
Db 420 YRYSLLLLGPKNPSGIILDLLELLELRLHQLWSPQAORNIISLQDICYAPLNPD 479
Qy 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTA 539
Qy 540 LSCMADYGAVPFPPLAIGYKGDYSEABALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 599
Db 540 LSCMADYGAVPFPPLAIGYKGDYSEABALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 599
Qy 600 EMRAFQRMAGMFOVTTAERSLEDEINRTTAEDLPIFATSVIVIFLYISALGSYSWS 659
Db 600 EMRAFQRMAGMFOVTTAERSLEDEINRTTAEDLPIFATSVIVIFLYISALGSYSWS 659
Qy 660 VMVDSKATLGLGGVAVVLGAWMAAGFFSYLGIRSSVLIVQVFPFLVLSVGADNIFIV 719
Db 660 VMVDSKATLGLGGVAVVLGAWMAAGFFSYLGIRSSVLIVQVFPFLVLSVGADNIFIV 719
Qy 720 EYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSG 779
Db 720 EYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSG 779


```
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-736-769-2

Query Match      78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLLRQAQSBPYTHIQPGYCAFYDECGKPELSGSLMTLSNVSCLSN 60
Db 1 MAEAWL-GWLLWALLLSAAGELYTPKHEAGVCTFYEBCKPELSGGTSLSNVSCLSN 59

QY 61 TPARKITGDHLILLOKICPRLYTGPNTQ-ACCSAKQLVLEASLSITKALLTRCPACSDN 119
Db 60 TPARKVTGEHALLOKICPRLYNGPNTTFACCSKQLLSLESSMSITKALLTRCPACSDN 119

QY 120 FYNLCHNTCSNOSLFINVTRVAQIAGQLPAVVAEYAFYQHSFAEQSYDCSRVRVPA 179
Db 120 FVSLCHNTCSNOSLFINVTRVAQIAGQLPAVVAEYAFYQHSFAEQSYDCSRVRVPA 179

QY 180 AATLAVGTMGVYGALCNAQRLWLFQDGTGNGLAPLDTFHLLEPGQAVGSGIOPANEK 239
Db 180 AASLAVGSMGVYGALCNAQRLWLFQDGTGNGLAPLDTFHLLEPGQAVGSGIOPANEK 239

QY 240 VARNESQDDVATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVPVVT 299
Db 240 IAPCHESQDDVATCSQDCAASCPVIPPPEALRSPFYNGRMPGSLVLIILCSVPVVT 299

QY 300 ILLVGRVAPARDKSMVDPKKGTSLSDKLSFTHTLLGQFGGWTWVASWPLTILVLS 359
Db 300 AVLRLRVVNRNKNKAQBPQAPKLPKHKLSPHTILGRPFQNGTRVASWPLTILVLS 359

QY 360 VTPVVALAAGLVFTLTDPVLEWAPNSQARSEKAFHDOHEGPPFRTNQVILTAPESS 419
Db 360 FIVVIALAAGLVFTLTDPVLEWAPNSQARSEKAFHDOHEGPPFRTNQVILTAPESS 419

QY 420 VRYDSLLGPKNFSGLDLDLLELLELQERHLQVWSPQAQRNLSQDICVAPLNPDN 479
Db 420 KYDYSLLGKNSGILSLDFLELLELQERHLQVWSPQAQRNLSQDICVAPLNPDN 479

QY 480 TSLYDCCNSLLOQYFQNNRTLLLTANQTLMGQTSQVDMKHFLYCANAPLTFKDGTLA 539
Db 480 TSLSDCCNSLLOQYFQNNRTLLLTANQTLMGQTSQVDMKHFLYCANAPLTFKDGTLA 539

QY 540 LSCMADYGAPVPPFLAIGYKGYDYSEBALIMTSLNYPAGDPRLAQAKLWEAFLEB 599
Db 540 LSCMADYGAPVPPFLAIGYKGYDYSEBALIMTSLNYPAGDPRLAQAKLWEAFLEB 599

QY 600 NRAFQRMAGMFOVTFATERSLEDEINRTTAEDLPFATSYIVIFLYTSLALGSYSSWR 659
Db 600 MESFORNTSDKQVAFSERSLEDEINRTTIQDLFPVAVSYIIVIFLYTSLALGSYSSWR 659

QY 660 VVWDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSVLQVVPFLVLSVGADNIFIFVL 719
Db 660 VAVESKATLGLGVAVVLGAVMAAMGFFSYLGPVSSVLVQVVPFLVAVGADNIFIFVL 719

QY 720 EYQRLPRRPGEPREHVGIRALGRVAPSMMLCSISEALCFPLGALTMPAVRTEALTSGLA 779
Db 720 EYQRLPRRPGEPREHVGIRALGRVAPSMMLCSISEALCFPLGALTMPAVRTEALTSGLA 779

QY 780 VILDFLLQMSAFVALISLDSKQEAERLDVCCVCPQELPPPGQSGGLLIGFPQKAYAPP 839
Db 780 IILDFLLQMTAFVALLSLDSKQEAERLDVCCVCPQELPPPGQSGGLLIGFPQKAYAPP 839

QY 840 LLHWITRGVLLFLFALFGVSLYSKCHISVGLDQELALPKDSYLLDYFLPLNRYFVGVAP 899
Db 840 LLHREIRPVVMLFLTLFGANLYLMCNINVLGDLQELALPKDSYLLDYFLPLNRYFVGVAP 899

QY 900 VYFVTTGLYNFSSAGMNAICSSAGCNFSTFKIQYATEPPEQSLATPASSWVDDFID 959
Db 900 VYFVTTSGFNFSSAGMNAICSSAGCNFSTFKIQYATEPPEQSLATPASSWVDDFID 959

QY 960 WLTP-SSCCRLYISGPNKDKFCPTVNSLNCILKNCMSITMGSVRPSVEQPHKYLPMFLND 1018
Db 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCLKNCNRTLGPVPTAQPHKYLPMFLND 1019

QY 1019 RNIIKCPKGLAAYSTSVNLTSQVLAASRPMAYHKPLKNSQDYTEALRAARELANITA 1078
Db 1020 PNRIKCPKGLAAYSTSVNLSSDQVIAQFMAVHKPLRNSQDYTEALRAARELANITA 1079

QY 1079 DLKVPKGTDPAPPEVPPYITINVPYQYLTILPEGLPMLSLCLVPTPAVSCILGLDLSG 1138
Db 1080 DLKVPKGTDPPEVPPYITINVPYQYLTILPEGLPMLSLCLVPTPAVSCILGLDLSG 1139

QY 1139 LNLLSIIMILVDTVGFMAWMDISYNAVSLINLVAAGMSVGFVSHITRSPAISTKPTWL 1198
Db 1140 LNLLSIIMILVDTTGLMAVWGISYNAVSLINLVAAGMSVGFVSHITRSPAISTKPTWL 1199

QY 1199 ERAKATITSMGSAVAGVAMTNLPGILVLGLAKAQILQIFPPRLNLLITLGLLHGLVFL 1258
Db 1200 ERAKATITVMSGSAVAGVAMTNFPGLITLGLFAQAQILQIFPPRLNLLITLGLLHGLVFL 1259

QY 1259 PVLSYVGVDPVNPALALRQKRAEAAVAVMVASCPNHRPSRVSTADNIYNHSPGSI-KG 1317
Db 1260 PVLSYVGVDPVNPALALRQKRAEAAVAVMVASCPNHRPSRVSTADNIYNHSPGSI-KG 1317

QY 1318 AGAISNFPNNGROF 1332
Db 1317 ANAARSSLPKSDQKP 1331

RESULT 15
US-10-621-758A-12
; Sequence 12, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-621-758A-12

Query Match      78.3%; Score 5407; DB 4; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
```

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
Qy	4	AGLRGMLLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA	63
Db	3	AAWQGMWALLLNSAQGLYTPHKGAGCTFYEECGKNPELSGGLTSLNSICLSNTPA	62
Qy	64	RKITGDHLLILQKICPRLYTGN - TOACCSAKQLVSLASISITKALLTRCPACSDNFVN	122
Db	63	RHVTGDHLLALQRCVPRLYNGENDTYACCTKQLVSLDSSLITKALLTRCPACSENFVS	122
Qy	123	LHCHNTCSNQSOLFNVTRVAQLGQOLPAVVAEAFYQHSFAQSVDSCSRVVPAAAT	182
Db	123	IHCNTCSPDOSLFNVTRVQRPDQQLPAVVAEAFYQHSFAKAYECSRVRIPAAAS	182
Qy	183	LAVTGTCGVSALCNAQRWLNFGQDTGNGLAPLDTITFHLLEPGQAVGSGIQPLNEGVAR	242
Db	183	LAVSMCGVYGALCNAQRWLNFGQDTGNGLAPLDTITFHLLEPGQALADGKPLDGKITP	242
Qy	243	CNESQGGDDVATSCQDCAAACPAIARPOALDSTFYLGOMPSGLVLIILCSVFAVVTILL	302
Db	243	CNESQGEDSAACSCQDCAAACCPVIPPALRPSFYMGPMGWLALIIIFTAVFLLSVVL	302
Qy	303	VQFRVAPARDKKNVDPKCTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSVP	362
Db	303	VYLRVASNRKNKTAGSQEAPNLPRKRRFSPHTVLGRFFESWGTRVASWPLTILVLSFV	362
Qy	363	VVALAAGLVFTELTPDPVELWAPNSQARSEKAFHDQHPGPFRTNQTILAPNRSYRY	422
Db	363	VIALSVGLTFIELTDPVELWAPNSQARSEKAFHDEHGFPPFRINQVFTAKNRSYKY	422
Qy	423	DSLLGPKNFSGLDLDLLELELERLHQLVMSPEAQRNISLQDICYAPLNPDNTSL	482
Db	423	DSLLGPKNFSGLSLLDQELLEQLERLHQLVMSHEAQRNISLQDICYAPLNPDNTSL	482
Qy	483	YDCCINSLLQYFQNNRTILLTANQTLMGQTSQVDMKQHFYCANAPLTFKDGTLALSC	542
Db	483	TDCCVNSLLQYFQNNHTLLLTANQTLNGQTSQVDMKQHFYCANAPLTYKDGTLALSC	542
Qy	543	MADYGAPVPFPLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEMPA	602
Db	543	IADYGAPVPFPLAVGYQGTDYSEAEALITFSINYPADDPMAHAKLWEEAFLKEMQS	602
Qy	603	FORMAGMFOVTTAERSLEDEINRTTADLPIFATSYIVIFLYISIALGSSYSNRVMV	662
Db	603	FORSTADKFQIAPSAERSLEDEINRTTIQDLPVFAISLYIVFLYISIALGSSYSNRVAV	662
Qy	663	DSKATILGGLGVAVVLGAVVAAMGFYSYLGVPSSLVIIQVVPFLVAVGADNIFVLEYQ	722
Db	663	DSKATILGGLGVAVVLGAVVAAMGFYSYLGVPSSLVIIQVVPFLVAVGADNIFVLEYQ	722
Qy	723	RLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALCPFLGALTMPAVRTPALTSGLAVIL	782
Db	723	RLPRMPGEQREAHIGRLGSAVPSMLLCSLSEALCPFLGALTMPAVRTPALTSGLAII	782
Qy	783	DFLQMSAFVALLSLDSKQREASRLDVCCVKQDELPPQCGEGILLIGFQKAYAPLLH	842
Db	783	DFLQMTAFVALLSLDSKQREASRPDVVCCFSERNLPPPKQKELLCPFRKIYTPPLH	842
Qy	843	WITRGVULLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVFGAPYV	902
Db	843	RFIRPVULLFLVLFGANLYUMCNSISVGLDQDLALPKDSYLLDYFLFLNRYLEVGPYV	902
Qy	903	VTLTGYNFSSBAGNNAICSSAGCENFPTQKIQVATEFPPEQSYLAIPASGWDDFDIWL	962
Db	903	DTTSGYNFTEAGNNAICSSAGCESFSLTKIQVASEFPNQSYYAIAASSWDDFDIWL	962
Qy	963	P-SSCCRYISGPNKDFCPSTVNSLNCIKMNSITMGSVRPSVEQHKYLPWFINDRPN	1021
Db	963	PSSSCCRYITRGPHKDFCPSTDTSFNCLKNMNRITGVPRTTQEQHKYLPWFINDTPN	1022
Qy	1022	IKCPKGGIAAYSTSVNLTSQVILASRFMAYHKPLKNSQDYTEALRAARELANITADLR	1081
Db	1023	IRCPKGGIAARTSVNLSSDGOI IASQFMAHYKPLRNSQDFTFALRASLLAANITAEIR	1082

Qy	1082	KVPGTDPAPFVRPYTTITNVFYEQYLITLPEGLPMLSLCLVPTFAVSCLLIAGLDIRSGLN	1141
Db	1083	KVPGTDPNFVFPYTTISNVFYQYLITLPEGLITFLALCFVPTFVVCVLLIAGLDIRSGLN	1142
Qy	1142	LLSIVMLLVDTVGFMALWDISYNAVSLINIVSAGMSVEFVSHITRSFAISTKPTWLEA	1201
Db	1143	LLSIIMLLVDTIGLMAVWGISYNAVSLINIVTAVGMSVEFVSHITRSFAVSTKPTRLERA	1202
Qy	1202	KEATISMGSAVFAVAMTNIPLGILVGLAKAQILQIOPFFRNLNLLITLLGLLHGLVFLPVI	1261
Db	1203	KDATIFMGSAVFAVAMTNPFGILILGFAQOLIQIOPFFRNLNLLITLLGLLHGLVFLPVI	1262
Qy	1262	LSYVGPDVNPALAEQKRAEAAVAVVAVASCPNHPRSVSTADNIIYNHSEGS- IKGAGA	1320
Db	1263	LSYUGPDVNPQALVLEKATEA-AMWSEPCQPPFPADANTSDYVNVGFNPFPIEINA	1321
Qy	1321	ISNPLPNNGRQF 1332	
Db	1322	ASSSLPKSDQKF 1333	
RESULT 16			
US-10-663-208A-12			
; Sequence 12, Application US/10663208A			
; Publication No. US20040132058A1			
; GENERAL INFORMATION: Scott W			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Wang, Luquan			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JB01603K2 US			
; CURRENT APPLICATION NUMBER: US/10/663,208A			
; PRIORITY FILING DATE: 2003-09-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; NUMBER OF SEQ ID NOS: 50			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 1333			
; TYPE: PR1			
; ORGANISM: Mus sp.			
US-10-663-208A-12			
Query Match 78.3%; Score 5407; DB 4; Length 1333;			
Best Local Similarity 77.2%; Pred. No. 0;			
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
Qy	4	AGLRGMLLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA	63
Db	3	AAWQGMWALLLNSAQGLYTPHKGAGCTFYEECGKNPELSGGLTSLNSICLSNTPA	62
Qy	64	RKITGDHLLILQKICPRLYTGN - TOACCSAKQLVSLASISITKALLTRCPACSDNFVN	122
Db	63	RHVTGDHLLALQRCVPRLYNGENDTYACCTKQLVSLDSSLITKALLTRCPACSENFVS	122
Qy	123	LHCHNTCSNQSOLFNVTRVAQLGQOLPAVVAEAFYQHSFAQSVDSCSRVVPAAAT	182
Db	123	IHCNTCSPDOSLFNVTRVQRPDQQLPAVVAEAFYQHSFAKAYECSRVRIPAAAS	182
Qy	183	LAVTGTCGVSALCNAQRWLNFGQDTGNGLAPLDTITFHLLEPGQAVGSGIQPLNEGVAR	242
Db	183	LAVSMCGVYGALCNAQRWLNFGQDTGNGLAPLDTITFHLLEPGQALADGKPLDGKITP	242
Qy	243	CNESQGGDDVATSCQDCAAACPAIARPOALDSTFYLGOMPSGLVLIILCSVFAVVTILL	302
Db	243	CNESQGEDSAACSCQDCAAACCPVIPPALRPSFYMGPMGWLALIIIFTAVFLLSVVL	302

```
QY 303 VGRVAPARDKSKWDPKKGTSLSDKLSPTSTHTLLGQFQCGWGTWVASHPLTILVLVSVP 362
Db 303 VYLRVASNRNKNKTAGSQEAPNLPKRFRPSHTVLGRFPESWGTVRVASHPLTILVLSFIV 362
QY 363 VVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGPGRPTNQVILLTAPNRSSYRY 422
Db 363 VIALSVGLTFTLTDPVELWSAPNSQARSEKAFHQHFGPGRPTNQVILLTAPNRSSYRY 422
QY 423 DSSLGPKNFGSGLDLDLLELELELERLRLHQLVMSPEAQNRNLSLQDICYAPLNPDNTSL 482
Db 423 DSSLGPKNFGSGLDLDLLELELELERLRLHQLVMSPEAQNRNLSLQDICYAPLNPDNTSL 482
QY 483 YDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAALASC 542
Db 483 TDCCVNSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAALASC 542
QY 543 MADYGAPVFPFLAIGYKGYKDYSEAEALJWTSFLNNYPAGDPDLAQAOKLWEEAFLEEMRA 602
Db 543 IADYGAPVFPFLAIGYKGYKDYSEAEALJWTSFLNNYPAGDPDLAQAOKLWEEAFLEEMRA 602
QY 603 FORMMAGMFOVTTFAERSLEDEINRTTAEIDLPIFATSYIVIFLYISLALGSYSSWRSVMV 662
Db 603 FORMMAGMFOVTTFAERSLEDEINRTTAEIDLPIFATSYIVIFLYISLALGSYSSWRSVMV 662
QY 663 DSKATILGCGVAVVILGAVMAAGFFSYLGRSLVILQVVPFLVLSVGADNIFIFVLEIYQ 722
Db 663 DSKATILGCGVAVVILGAVMAAGFFSYLGRSLVILQVVPFLVLSVGADNIFIFVLEIYQ 722
QY 723 RLPRRPGEBREHIGALGRVAPSMLLCSLSAICFFLALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRRPGEBREHIGALGRVAPSMLLCSLSAICFFLALTPMPAVRTFALTSGLAIVL 782
QY 783 DFLQMSAFVALLSDSKRQESRLDVCCKVPQELPPGQEGLLGLFFQKAYAPFLILH 842
Db 783 DFLQMSAFVALLSDSKRQESRLDVCCKVPQELPPGQEGLLGLFFQKAYAPFLILH 842
QY 843 WITRGVLLFLFALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYEFGVAPYF 902
Db 843 WITRGVLLFLFALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYEFGVAPYF 902
QY 903 VTTLGVSFSEAGMAICSSAGCNFSFTOKIYATEPPEQSVAIPASSWDDDFIDWLT 962
Db 903 VTTLGVSFSEAGMAICSSAGCNFSFTOKIYATEPPEQSVAIPASSWDDDFIDWLT 962
QY 963 P-SSCCRLYISGPNKDFCPTVNSLNCILKNQMSITMGSVRPSVEQFHKYLPWFNDNRN 1021
Db 963 P-SSCCRLYISGPNKDFCPTVNSLNCILKNQMSITMGSVRPSVEQFHKYLPWFNDNRN 1021
QY 1022 IKCPKGLAAYSTSVNLTSQGVLASRFMAHKLKNSQDYTEALRAARELANITADLR 1081
Db 1022 IKCPKGLAAYSTSVNLTSQGVLASRFMAHKLKNSQDYTEALRAARELANITADLR 1081
QY 1082 KYPGTDPAPEVFPYITVNFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLIN 1141
Db 1082 KYPGTDPAPEVFPYITVNFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLIN 1141
QY 1142 LLSIVMLVDVTFWMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1201
Db 1142 LLSIVMLVDVTFWMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1201
QY 1202 KEATISMGSAVPAVAMTNLPGLVILGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1261
Db 1202 KEATISMGSAVPAVAMTNLPGLVILGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1261
QY 1262 LSVYGPVNPALALQKRAEBAVAVMVASCPNHFSRSTADNITYNHSFEGS-1KGAGA 1320
Db 1262 LSVYGPVNPALALQKRAEBAVAVMVASCPNHFSRSTADNITYNHSFEGS-1KGAGA 1320
QY 1321 ISNFLPNNGROF 1332
Db 1322 ASSSLPKSDQKF 1333
```

```
RESULT 17
US-10-646-301A-12
; Sequence 12, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-646-301A-12
```

Query Match 78.3%; Score 5407; DB 4; Length 1333;

Best Local Similarity 77.2%; Pred. No. 0;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

```
QY 4 AGLRGWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELGSIMTISNVLSCLNTPA 63
Db 3 AAWQGWLLWALLRLNSAQSELYPTHKAGCTFYDECGKNPELGSIMTISNVLSCLNTPA 62
QY 64 RKTGDHLLILQKICPLRYTGN-TOACCSAKQLVSLASLSTKALLTRCPACSNVYN 122
Db 63 RHVTGDHLLALQRCVCPRLYNGPNDTYACCSTKQLVSLSSLSITKALLTRCPACSNFVS 122
QY 123 LHCHNTCPNQSLFNVTNRVAQLGAGQLPVAVYEAIFYQHSFAEOSYDSCSRVRVPAAT 182
Db 123 IHCHNTCPDQSLFNVTNRVQDQGLPQVAVYEAIFYQHSFAEKAYESCSRVRIPAAAS 182
QY 183 LAVGTMCVGYGALCNAQRWLNFGQDTGNGLAPLBITFHLBPQAVGSGIQLNPGVAR 242
Db 183 LAVGSMCVGYGALCNAQRWLNFGQDTGNGLAPLBITFHLBPQAVGSGIQLNPGKITP 242
QY 243 CNESQGDVATCSCDCAASCPAIARPQALDSTFVYGMPGSLVLIILCSFVAVVTILL 302
Db 243 CNESQGBSAAASCQDCAASCEVIPPALRPSFTYMGMPGWLIIIFAVFVLLSVVL 302
QY 303 VGRVAPARDKSKWDPKKGTSLSDKLSPTSTHTLLGQFQCGWGTWVASHPLTILVLVSVP 362
Db 303 VYLRVASNRNKNKTAGSQEAPNLPKRFRPSHTVLGRFPESWGTVRVASHPLTILVLSFIV 362
QY 363 VVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGPGRPTNQVILLTAPNRSSYRY 422
Db 363 VIALSVGLTFTLTDPVELWSAPNSQARSEKAFHQHFGPGRPTNQVILLTAPNRSSYRY 422
QY 423 DSSLGPKNFGSGLDLDLLELELELERLRLHQLVMSPEAQNRNLSLQDICYAPLNPDNTSL 482
Db 423 DSSLGPKNFGSGLDLDLLELELELERLRLHQLVMSPEAQNRNLSLQDICYAPLNPDNTSL 482
QY 483 YDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAALASC 542
Db 483 TDCCVNSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAALASC 542
QY 543 MADYGAPVFPFLAIGYKGYKDYSEAEALJWTSFLNNYPAGDPDLAQAOKLWEEAFLEEMRA 602
Db 543 IADYGAPVFPFLAIGYKGYKDYSEAEALJWTSFLNNYPAGDPDLAQAOKLWEEAFLEEMRA 602
QY 603 FORMMAGMFOVTTFAERSLEDEINRTTAEIDLPIFATSYIVIFLYISLALGSYSSWRSVMV 662
Db 603 FORMMAGMFOVTTFAERSLEDEINRTTAEIDLPIFATSYIVIFLYISLALGSYSSWRSVMV 662
```

QY	663	DSKATLGLGGVAVVLGAVMAAMGFFSYGIRSSLVILQVVPFLVLSVGCADNIFIFVLEYQ	722
Db	663	DSKATLGLGGVAVVLGAVMAAMGFFSYGLVFPSSLVIIQVVPFLVAVGADNIFIFVLEYQ	722
QY	723	RLPRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLGALTTPMPAVRTFALTSGLAVAL	782
Db	723	RLPRMPGQREAHIGRTLGSVAPSMMLCSLSEAICFFLGALTTPMPAVRTFALTSGLAIF	782
QY	783	DFLLQMSAFVALLSLDSKQREASRLDVCCVKPQELPPPGQEGLLGLFFQKAYAPFLH	842
Db	783	DFLLQMTAFVALLSLDSKQREASRDVCCVSSRNLPKPKQEGLLGLFFKAYVTPFLH	842
QY	843	WTRGVVLLFLALFGVSLYSWCHISVGLDQELALPKDSYLLDIFLFLNRYFEVGPVYF	902
Db	843	REIRPVVLLFLVLFGANLYLMCNISVGLDQDALPKDSYLLDIFLFLNRYLEVGPVYF	902
QY	903	VTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT	962
Db	903	DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIQYASEFPNQSYVAIAASSWVDDFIDWLT	962
QY	963	P-SSCCRLYISGPNKDFCPTVNSLNCILKNCMSITMGSVRPSVQPHKYLPMFLNDRPN	1021
Db	963	PSSSCCRIYTRGPHKDFCPTDTSFNCLKNCMNRTLGVRPTTBQPHKYLPMFLNDRPN	1022
QY	1022	IKCPKGLLAAYSTVNLSGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLR	1081
Db	1023	IRCPKGLLAAYSTVNLSGQIIAASQFMAYHKPLRNSQDFTALRASRLAANITAEUR	1082
QY	1082	KVPGTDPAPFVPPYITINVFYEQYLTILPEGLPMLSLCLVPTFAVSCLLGLDLRSGLIN	1141
Db	1083	KVPGTDPNFEVPPYITINSVFYQYLTIVLEGIFTALCFVTFVVCYLLGLDIRSGLIN	1142
QY	1142	LLSIVMILVDITGFMALMDSYNVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER	1201
Db	1143	LLSIIMLVDTIGLMAVWGISYNVSLINLVTAVGMSVEFVSHITRSPAVSTKPTRLERA	1202
QY	1202	KENTISMGSAVPAVAVMTNLPGLIVLGLAKAQLIQIFFERLNLITLGLLHGLVLPVI	1261
Db	1203	KOATIFMGSAVPAVAVMTNFPGLIILGFAQAQLIQIFFFLNLLITLGLLHGLVLPV	1262
QY	1262	LSVVGDPVPALALEOKRAEEAAVAVMVASCPNHPSRVSTADNIYVNHSEFSGS-1K320	1320
Db	1263	LSVVGDPVQVAVLEKLALEA-AMVSEPCQYPPADANTSDYVNYGFNPEFPEINA	1321
QY	1321	ISNFLPNNGRQF 1332	
Db	1322	ASSSLPKSDQKF 1333	
RESULT 18			
US-10-736-769-12			
; Sequence 12, Application US/10736769			
; Publication No. US20040161838A1			
; GENERAL INFORMATION:			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Wang, Luquan			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JB01603-K3-US			
; CURRENT APPLICATION NUMBER: US/10/736,769			
; CURRENT FILING DATE: 2003-12-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; PRIOR APPLICATION NUMBER: 10/663,208			
; PRIOR FILING DATE: 2003-09-16			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 1333			
; TYPE: PRT			
; ORGANISM: Mus sp.			
US-10-736-769-12			
Query Match			
Best Local Similarity 78.3%; Score 5407; DB 4; Length 1333;			
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
QY	4	AGURGLLWALLLRLAQSEBPTTHQGYCAFYDECKNPBLSGLTSLTNVSCLSNTPA	63
Db	3	AAMQGLLWALLLSAQGELYTPTHKAGCTTFYBECCKNPBLSGLTSLTNVSCLSNTPA	62
QY	64	RKITGDHLILLOKICPLRYLTGPN--TOACCSAKOLVLSLEASISITKALLITRCPACSDNFVN	122
Db	63	RHVTGDHLLALQVRCPRLYNGPNDTYACCSTKQLVLSLSSLSITKALLITRCPACSENFVS	122
QY	123	LHCNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFABOSYDSCRVRVPAAT	182
Db	123	IHCNTCSPDQSLFINVTRVVRDPGQLPAVVAYEAFYQSFABKAYECSRVRIPAAAS	182
QY	183	LAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITPHLLEPGOAVSGIQLNEGVAR	242
Db	183	LAVSGMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITPHLLEPGOALADGKPLDKTTP	242
QY	243	CNESQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIITLCSFVAVVTILL	302
Db	243	CNESQGEDSAACSQDCAASCVPVPPPALRPSFYMGRMPCWALIIITFTAVFVLLSVVL	302
QY	303	VGFRVAPARDKSMVDPKKJLSDKJSFSTHTLLGQFQGWGTWVASWPLTILVLSVTP	362
Db	303	VYLRVASNRKNKTAGSQEAPNLPRKRFPSPHTVLGRFFESWGTRVASWPLTVLALSIV	362
QY	363	VVALAAGLVFTELTTDPVELASAPNSQARSKAFHQHFGFPFRTNQVILTAPNRSSYRY	422
Db	363	VIALSVGTIFELTTDPVELWSAPKSQARKKAFHDEHFGFPFRINQVITAKNRSSYKY	422
QY	423	DSLILGPKNFSGILDLLELLELQERLRLHQLVMSPEAQRNLSLQDICYAPLNPDTSL	482
Db	423	DSLILGPKNFSGILSLDLLELLELQERLRLHQLVMSHEAQRNLSLQDICYAPLNPHTSL	482
QY	483	YDCCINSLLQVFPNNRTRLLLTANQTLMGQTSQVDMKQHFLYCANAPLTKDGTALALSC	542
Db	483	TDCCVNSLLQVFPNNHTLLLTANQTLNGQTSQVDMKQHFLYCANAPLTKDGTALALSC	542
QY	543	MADYCAPVPPFLAIGYKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEEAFLEEMRA	602
Db	543	IADYCAPVPPFLAVGGYQGTDIYSEAEALITFESINNYPADDDPRMAHAKLWEEAFLEKMQS	602
QY	603	FORRMAGMFQVTFRTAERSLEDEINRTTABDLPIFATSYIVIFLYTISLALGSYSSSRVMV	662
Db	603	FORSTADKQFIAPSAERSLEDEINRTTIQDLVPFAISYLVIFLYTISLALGSYSSSRVAV	662
QY	663	DSKATLGLGGVAVVLGAVMAAMGFFSYGIRSSLVILQVVPFLVLSVGCADNIFIFVLEYQ	722
Db	663	DSKATLGLGGVAVVLGAVMAAMGFFSYLQIRSSLVILQVVPFLVLSVGCADNIFIFVLEYQ	722
QY	723	RLPRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLGALTTPMPAVRTFALTSGLAVAL	782
Db	723	RLPRMPGQREAHIGRTLGSVAPSMMLCSLSEAICFFLGALTTPMPAVRTFALTSGLAIF	782
QY	783	DFLLQMSAFVALLSLDSKQREASRLDVCCVKPQELPPPGQEGLLGLFFQKAYAPFLH	842
Db	783	DFLLQMTAFVALLSLDSKQREASRPDVCCVSSRNLPKPKQEGLLGLFFKAYITVTPFLH	842
QY	843	WTRGVVLLFLALFGVSLYSWCHISVGLDQELALPKDSYLLDIFLFLNRYFEVGPVYF	902
Db	843	REIRPVVLLFLVLFGANLYLMCNISVGLDQDALPKDSYLLDIFLFLNRYLEVGPVYF	902
QY	903	VTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT	962
Db	903	DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIQYASEFPNQSYVAIAASSWVDDFIDWLT	962

QY 963 P-SSCCRLYSIGNKDKFCBPTVNSLNCIKNCMSITMGSVRPSVEQPHKYLPHFLNDREN 1021
Db 963 PSSCCRLYTRGPHKDEFCBPTDTSFNCCLKNCMNRTGLPVRPTTEQPHKYLPHFLNDTEN 1022
QY 1022 IKCPKGLAAYTSVNLTSQGVLAGRFMAYHKPLKNSODYTEALRAARELAANITADLR 1081
Db 1023 IRCPKGLAAYTSVNLSSDGOIIASQFMAHYKPLNSQDFTALRASRLAANITABLR 1082
QY 1082 KVPGTDPAPFVPPYTTTNNVPEYQYLTILPEGLPMLSLCVPTPAVSCLLLGLDLRSLN 1141
Db 1083 KVPGTDPENFVPPYTTTNSVYQYLTALPEGITLALCFVPTVVCYLLGLDIRSGILN 1142
QY 1142 LLSIVMLVDTYGFMAWISYNAVSLINLVSAGVMSVEFVSHITSPALSTKPTWLER 1201
Db 1143 LLSIIMLVDTIGLMAVGLSYNAVSLINLVTAVGMSVEFVSHITSPAVSTKPTWLER 1202
QY 1202 KBATISMGSAVAGVANTMLPGILVLGLAKAQLIQIFFPRLNLLITLLGLLHGLVLPVI 1261
Db 1203 KBATIFWGSNAVAGVANTWFPGLIILGPAQAQLIQIFFPRLNLLITLLGLLHGLVLPV 1262
QY 1262 LSYVGPDPNPALALEOKRAEBAVAAVMVASCPNHPRSVSTADNIVYNHSFEGS- IKGAGA 1320
Db 1263 LSYLGPDPVNAQVLEEKATEA-AMVSEPCQYPPPADANTSDYVNYGPNPEFIPINA 1321
QY 1321 ISNFLPNRQF 1332
Db 1322 ASSSLPKSQKF 1333

RESULT 19

US-10-450-763-53050
; Sequence 53050, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53050
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53050

Query Match 64.6%; Score 4466; DB 5; Length 982;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 867; Conservative 1; Mismatches 9; Indels 8; Gaps 3;
QY 1 MAEAGLRGWLWALLRLAQSEPYTTTHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTTHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLLOKICPRLTYGNTQACCSAKQLVLEASLITKALLTRCPACSNDF 120
Db 61 TPARKITGDHLLLOKICPRLTYGNTQACCSAKQLVLEASLITKALLTRCPACSNDF 120
QY 121 VNLHCHNTCSNQSIFINVTVAQLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSIFINVTVAQLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180

QY 181 ATLAVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVGSGIQPLNEG 240
Db 181 ATLAVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVGSGIQPLNEG 240
QY 241 ARCNSQGDVATCSQDCAASCP--AIARPOALDSTFYLGQMPGSLVLIILILCSVAV 238
Db 241 ARCNSQGDVATCSQDCAASCP--AIARPOALDSTFYLGQMPGSLVLIILILCSVAV 238
QY 299 TILLVGRVAPARDKSKVDPKKGTSLSDKLSFSTHTLLGQPFQGTGWTWASWPLTILVL 358
Db 299 TILLVGRVAPARDKSKVDPKKGTSLSDKLSFSTHTLLGQPFQGTGWTWASWPLTILVL 358
QY 359 SVIPVVALAAGLVFTTELTTDPVELMSAPNSQARSEKAFHDQHPFPFRNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTELTTDPVELMSAPNSQARSEKAFHDQHPFPFRNQVILTAPNRS 418
QY 419 SYRYSLLGPKNFGSGLDLDLLELLEQLERHLQVMSPEAQRNLSLQDICYAFLNPD 478
Db 419 SYRYSLLGPKNFGSGLDLDLLELLEQLERHLQVMSPEAQRNLSLQDICYAFLNPD 478
QY 479 NTSLYDCCINSLLQVFQNNRTLLLTANOTLMGQTSQVDMKDHFLYCANAPLTFKDG 538
Db 479 NTSLYDCCINSLLQVFQNNRTLLLTANOTLMGQTSQVDMKDHFLYCANAPLTFKDG 538
QY 539 ALSCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLE 598
Db 539 ALSCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLE 598
QY 599 ENRAFQRMAGMFOVTFPFAERSLEDEINRTTAEDLPFATSYVIVPLYTSLAGSYSSWS 658
Db 599 ENRAFQRMAGMFOVTFPFAERSLEDEINRTTAEDLPFATSYVIVPLYTSLAGSYSSWS 658
QY 659 RVWVDSKATLGLGGVAVVLGAVMAAGPSPYIGIRSSLVILQVWPVLVSVGADNIFIFV 718
Db 659 RVWVDSKATLGLGGVAVVLGAVMAAGPSPYIGIRSSLVILQVWPVLVSVGADNIFIFV 718
QY 719 LEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTPMPAVRTFALTSG 778
Db 719 LEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTPMPAVRTFALTSG 778
QY 779 AVILDFLOMSAFVALLSDSKROBASRLDVCCCKVPQBELPPPGQEGILLGFFQKAYAP 838
Db 779 AVILDFLOMSAFVALLSDSKROBASRLDVCCCKVPQBELPPPGQEGILLGFFQKAYAP 838
QY 839 FLHWHITRGW---LLLFLALFGVSLYSYMSCHISVGLDQELALPK 879
Db 839 FLHWHITRGWVPSQLLLFLALFGVSLYSYMSCHISVGLDQELALPK 883

RESULT 20

US-10-208-731-2
; Sequence 2, Application US/10208731
; Publication No. US20030092038A1
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/10/208,731
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-208-731-2


```
Query Match      34.8%; Score 2402.5; DB 4; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.2e-199;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLAQSEPTTIIHQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVWYGEGLIAYGDKRYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGENTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKDLQLPLQFLSRCPSCFYN 103
QY 120 FVNLHCHNTCSNQSFLINVT-----VAQLGAGQLPAVVAEAFYQHSFAQSQSDSCSRV 175
DB 104 LNLFLCELTCSPQSQFLNVTATEDYDVPVNTQTKNVKELQYVYGQSFANAMYNACRDV 163
QY 176 RYPAATLAVGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDIT-----PHLEPGQA 228
DB 164 EAPSSNDKALGLLCKGKADA-CNATWIEYMFKNKGQAPFTITPVFSDPFVH-----215
QY 229 VSGGIQPLNEGVARCNESQDDVATCSCQDCASCPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTPACSCQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSKWDPKKGTSLS 326
DB 273 WITYMAFLVFFGAFFAVMCRYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 229 VSGGIQPLNEGVARCNESQDDVATCSCQDCASCPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTPACSCQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSKWDPKKGTSLS 326
DB 273 WITYMAFLVFFGAFFAVMCRYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFQCGWGTWASWPLTILVSVIPVVALAAGLVFTTELTDPVELWSAP 386
DB 325 DPVSAAFEGCLRLFRWGSFCVRNPGCVTFPSLVFTACSSGLVFRVVTNPDVLSWAP 384
QY 387 NSQARSEKAFDHQGFPPFTNTQVILTAPNRSRYRDSLLGPK-NFSGILDLDLLELL 445
DB 385 SSQARLEKEYFOHGFPPFTTEQLIIRAPLTDKHIQYPYSGADVPFGPPLDIQLHQLV 444
QY 446 ELQERLHLQWSPAEARNISLODICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 445 DLQIAIEN--ITASYNDETVTLDICLAPLSPYNT---NCTILSVLNYFQNSHSLDHKK 499
QY 506 NQTLMGQTSQVDKDHFLICANAPLTFKGTALALSCMADYGAQVFPFELAIQYKGYKQDS 565
DB 500 GDDFF---VYADYHTFLYCVRAPASLNDTSLHDPLGTGFGFVPFWLVGLGYDQNTN 556
QY 566 EABALIMTFSLNYPAGDPRLAQKLWEEAFLEEMRAFORRMAGMFQVTFTAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTAEDLPATSYIVIFLYISALGYSYSSRWVDSKATILGGLGAVAVLGAVMAMG 685
DB 614 NRESDSDFVTWISYAIMFLYISLALGHKISCRRLAVDSKVSGLIAGILVLSSVACSGLG 673
QY 686 FFSVLGTRSSVILQVVPFLVLSVGADNIFIVLEVQLRPRRPGEPREHVHIGRALGRVAP 745
DB 674 VFSYIGLPLTILVIEVIFLVLAVGVNDIIFILVQAVORDERLOGETLDQQLGRVLGEVAP 733
QY 746 SMLLCSLSEACFPLGALTTPAVRTPALATSGLAVIDLDFLLQMSAFVALLSLDSKQGEAS 805
DB 734 SMFLSSFSSETVAFGLGALSNVPAVHTSLFAGLAVFIDFLQITCFVSLILGLDKRQKN 793
QY 806 RLDVCCCKPKQELPPPGQ-GEGLLLGFFQKAYAPFLHWTIRGVVLLFLALFGVSLYSM 864
DB 794 RLDIFCCVRGAEDGTSVQASESCULFRFPKNSYSPLLKDWNRPIVIAIFGVLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSVLLDYFLFLNRYFEVGAQVYFVTLGVNFSSEAGMAICSSAG 924
DB 854 NKVDIGLQSLMSPPDSTMYVDYFKSIQYLLHAGPPVYFVLEEGHDYTSKQNNVCCGMG 913
QY 925 CNNSFSTQKIQYATFEFQESYLAIPASSWVDDFDLWLTTP-SSCCRLYISGFNKDKFCPEST 983
DB 914 CNNSDLVQOIFNAQLDNYTRIGAPSSWIDDDYFDWKVQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNCSMIT-MGSVRSVRSQFHKYLFWFLNDRPNIKCPKGLLAAYSTSVN--LTS 1040
```

```
DB 971 VUDPACVR-CRPLTPEGKQRFQGDWRFPLPMFLSDNPNPKCGKGHAAVSSAVNILLGH 1029
QY 1041 DGQVLASRFMAYKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVFPVYITNV 1100
DB 1030 GTRVGATYFTYHTVLQTSADFIDALKCARLIASNVT-ETWINGS--AYRVFYSPYV 1086
QY 1101 FYEQYLTILPEGLPMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFMALWD 1160
DB 1087 FYEQYLTIIDTTFNLGVSLGALFVTLVWLLGCELWSAVINCATIAMVLVNMFGVWMLWG 1146
QY 1161 ISYNAVSLINLVSAGMSVEFVSHITRSPFAISTRTWLERAKEATISMGSAVFAQVAMTN 1220
DB 1147 ISLNAVSLVNLVMSGISVEFCSHITRAFTVSMKGSVERAEALAHMSSVFSGITLTK 1206
QY 1221 LPGAIVLGLAKAQILQIFFFLNLLIITLGLLHGLVFLPVILSYVGPVDNPA 1272
DB 1207 FGGIWLAFAKSQIFQIFYFRMYLAMVLLGATHGLIFLPLVILLSYIGPSVKA 1258

RESULT 21
US-10-741-601-530
; Sequence 530, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOISIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 530
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-530

Query Match      34.8%; Score 2402.5; DB 4; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.2e-199;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLAQSEPTTIIHQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVWYGEGLIAYGDKRYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGENTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKDLQLPLQFLSRCPSCFYN 103
QY 120 FVNLHCHNTCSNQSFLINVT-----VAQLGAGQLPAVVAEAFYQHSFAQSQSDSCSRV 175
DB 104 LNLFLCELTCSPQSQFLNVTATEDYDVPVNTQTKNVKELQYVYGQSFANAMYNACRDV 163
QY 176 RYPAATLAVGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDIT-----PHLEPGQA 228
DB 164 EAPSSNDKALGLLCKGKADA-CNATWIEYMFKNKGQAPFTITPVFSDPFVH-----215
QY 229 VSGGIQPLNEGVARCNESQDDVATCSCQDCASCPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTPACSCQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSKWDPKKGTSLS 326
DB 273 WITYMAFLVFFGAFFAVMCRYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFQCGWGTWASWPLTILVSVIPVVALAAGLVFTTELTDPVELWSAP 386
DB 325 DPVSAAFEGCLRLFRWGSFCVRNPGCVTFPSLVFTACSSGLVFRVVTNPDVLSWAP 384
QY 387 NSQARSEKAFDHQGFPPFTNTQVILTAPNRSRYRDSLLGPK-NFSGILDLDLLELL 445
DB 385 SSQARLEKEYFOHGFPPFTTEQLIIRAPLTDKHIQYPYSGADVPFGPPLDIQLHQLV 444
QY 446 ELQERLHLQWSPAEARNISLODICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 445 DLQIAIEN--ITASYNDETVTLDICLAPLSPYNT---NCTILSVLNYFQNSHSLDHKK 499
QY 506 NQTLMGQTSQVDKDHFLICANAPLTFKGTALALSCMADYGAQVFPFELAIQYKGYKQDS 565
DB 500 GDDFF---VYADYHTFLYCVRAPASLNDTSLHDPLGTGFGFVPFWLVGLGYDQNTN 556
QY 566 EABALIMTFSLNYPAGDPRLAQKLWEEAFLEEMRAFORRMAGMFQVTFTAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTAEDLPATSYIVIFLYISALGYSYSSRWVDSKATILGGLGAVAVLGAVMAMG 685
DB 614 NRESDSDFVTWISYAIMFLYISLALGHKISCRRLAVDSKVSGLIAGILVLSSVACSGLG 673
QY 686 FFSVLGTRSSVILQVVPFLVLSVGADNIFIVLEVQLRPRRPGEPREHVHIGRALGRVAP 745
DB 674 VFSYIGLPLTILVIEVIFLVLAVGVNDIIFILVQAVORDERLOGETLDQQLGRVLGEVAP 733
QY 746 SMLLCSLSEACFPLGALTTPAVRTPALATSGLAVIDLDFLLQMSAFVALLSLDSKQGEAS 805
DB 734 SMFLSSFSSETVAFGLGALSNVPAVHTSLFAGLAVFIDFLQITCFVSLILGLDKRQKN 793
QY 806 RLDVCCCKPKQELPPPGQ-GEGLLLGFFQKAYAPFLHWTIRGVVLLFLALFGVSLYSM 864
DB 794 RLDIFCCVRGAEDGTSVQASESCULFRFPKNSYSPLLKDWNRPIVIAIFGVLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSVLLDYFLFLNRYFEVGAQVYFVTLGVNFSSEAGMAICSSAG 924
DB 854 NKVDIGLQSLMSPPDSTMYVDYFKSIQYLLHAGPPVYFVLEEGHDYTSKQNNVCCGMG 913
QY 925 CNNSFSTQKIQYATFEFQESYLAIPASSWVDDFDLWLTTP-SSCCRLYISGFNKDKFCPEST 983
DB 914 CNNSDLVQOIFNAQLDNYTRIGAPSSWIDDDYFDWKVQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNCSMIT-MGSVRSVRSQFHKYLFWFLNDRPNIKCPKGLLAAYSTSVN--LTS 1040
```



```
Db 385 SSOARLEKEYFDHGFQFFRTEQLIIRAPLTKHIIQYIPSPGADVPFGPDLIOIHLQVL 444
Qy 446 ELQERLRHLQVMSPEARNI SLQDICYAPLNDNTSLYDCCINSLLIQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTIQDICLAPLSPNT---NCTILSVLNYFQNSHVLDDHKK 499
Qy 506 NQTLMGQTSQVDKDHFLICANAPLTFKDGITALALS CMADYCAPVFPFLAIGYKGYKQYS 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLIHDCLGTGCGPVPFVPLVGLGYDDQNYN 556
Qy 566 EBAELIMTSLNYPAGDPRLAQAQKLEAFLEEMAFQRRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAKEKEFINFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSYIVIFLYISLALGSYSMSRVMDSKATLGLGGVAVVLGAVMAANG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHIKSCRRLLVDSKVSIGIAGIILVSSVACSIG 673
Qy 686 FFSYLGIRSSLVILQVVPFLVSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQYQDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSAICFFELGALTPMPAVRTFALTSGTGLAVLDFLLQMSAFVALLSLSKQRAS 805
Db 734 SMLFSFSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIKRQKN 793
Qy 806 RLDVCCCKVQBLPPPGQ--GEGLLIGFFQKAYAPFLHWTIRGVVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASSELFRPFKNSYSPILLKDWMPRIPIAIFVGVLSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFNRYFEVGA PVYFVTTLGYNFSSBAGNNAICSSAG 924
Db 854 NKVDIGLQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEGRHDYTSKQGNMVCWGNG 913
Qy 925 CNPFSTQKIQVATEPEQSYLAI PASSWDDFDWLTP--SSCRLYISGPNKDFCPS 983
Db 914 CNDSLVQOIFNAALDNTYRIGFAPSSWIDDYFDMVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSMLCLNKMCSIT--MGSVRPSVEQFKYLPWFLNDRPNKCPKGLAAYSTSVN--LTS 1040
Db 971 VDPACVR--CRPLTEGKQRPQGGDMRFLPFLSDNPNPKCKGKHAAYSAAVAILLGH 1029
Qy 1041 DQGVLASRPMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNV 1100
Db 1030 GTRVGATFMTYHTVLOTSAFDALKKARLIASNT--ETWINGS--AYRVEPYSVFV 1086
Qy 1101 FYEQVLTILPEGLFMLSCLVTFPAVSCILGLDLSRGLNLLSIVMLIVDTGVPMALWD 1160
Db 1087 FYEQVLTIDDTIFNLGSLGAIFLVTVMVLLGCELWSAVIMCATIAMVLVNMFGYMWLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRGFAISTKPTMLERAKEATISMGSAVPAVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGIVSEFCSHITRAFTVSMKGSVEREAEALAHMGSSVFSGITLK 1206
Qy 1221 LPGAIVLGLAKAQLOIIPFRNLNLITLLGLHLVFLVPLVSYGVPDNP 1272
Db 1207 FGGIVVLAPAKSQIFQIFVFRMYLAWVLGATHGLIFLPLVLLSYIGSPSNKA 1258

RESULT 22
US-10-741-600-1542
; Sequence 1542, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1542
```

; LENGTH: 1278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-741-600-1542

Query Match 34.8%; Score 2402.5; DB 5; Length 1278;

Best Local Similarity 39.8%; Pred. No. 2.2e-199;

Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGLMLWALLLLRLAQSEPTTTHQPGVCAPYDRSG-----KNPELSGSLMTLSNVCSLS 59

Db 4 RGLALGLLLLLLCPAQVFSQ-----SCWYGECCGIAYGDKRYNCBSYG-----46

Qy 60 NTPARKITGDHLILQIKCPRLYTGTNTQACCSAQSLVLEASLSITTKALLTRCPACSDN 119

Db 47 ---PPRPLPKDGYDLVQELCPGFFG-NVSLCCDVQLQTLKNDLQLPLQFLSRCRSCFPN 103

Qy 120 FVNLCHNTCSNQSLFINVTR-----VAQLGAGQILPAVVAYEAFYQHSHFASQSYSDCSRV 175

Db 104 LNLFLCETCSPROQFLNVTATEDYVDPVNTQTKTNKELQYVVGQSPANAMYNACRDV 163

Qy 176 RVPAAATLAVGTMCGVYGSALCNAQRLNFOGDTGNGLAPLDIT-----PHLLEPGA 228

Db 164 EAPSSNDKALGLLCKGDADA--CNATNWIETMFKNDGQAPFTTITPVFSDFPVH-----215

Qy 229 VSGIQPLNEGVARCNESQGDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279

Db 216 ---GMEPMNNATKGDSEVDVATPCSDSIVCGPKPQPPPPPPAPWTILGLDAMYIM 272

Qy 280 QMPSGLVLIILCSVFAVVTILL-----VGRVAPARDKSMVDPKKGTSL 326

Db 273 WITTMALFVFGAFVAVMKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324

Qy 327 DKLSESTHILCQPPQGHGTWASWPLTILVSLVPPVVALAAGLVTELTTPDVELWSAP 386

Db 325 DPVSAAFEGCLRLRFTRWGSCVVRNPGCVIPFSLVFIACSSGLFVFRVTNTPVDLWSAP 384

Qy 387 NSQARSEKAFIDHQHGFPPFRNQVTLTAPNRSSRYDLSLLGPK--NFSGILLDLLELL 445

Db 385 SSOARLEKEYFDHGFQFFRTEQLIIRAPLTKHIIQYIPSPGADVPFGPDLIOIHLQVL 444

Qy 446 ELQERLRHLQVMSPEARNI SLQDICYAPLNDNTSLYDCCINSLLIQYFQNNRTLLLLTA 505

Db 445 DLQIAIEN--ITASVDNETVTIQDICLAPLSPNT---NCTILSVLNYFQNSHVLDDHKK 499

Qy 506 NQTLMGQTSQVDKDHFLICANAPLTFKDGITALALS CMADYCAPVFPFLAIGYKGYKQYS 565

Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLIHDCLGTGCGPVPFVPLVGLGYDDQNYN 556

Qy 566 EBAELIMTSLNYPAGDPRLAQAQKLEAFLEEMAFQRRMAGMFQVTFTAERSLEDEI 625

Db 557 NATALVITFPVNNYNDTEKQRAQAKEKEFINFVKNYKN---PNLTISFTAERSIEDEL 613

Qy 626 NRTTAEDLPFATSYIVIFLYISLALGSYSMSRVMDSKATLGLGGVAVVLGAVMAANG 685

Db 614 NRESDSVFTVVISYAIMFLYISLALGHIKSCRRLLVDSKVSIGIAGIILVSSVACSIG 673

Qy 686 FFSYLGIRSSLVILQVVPFLVSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAP 745

Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQYQDERLQGETLDQQLGRVLGEVAP 733

Qy 746 SMLLCSLSAICFFELGALTPMPAVRTFALTSGTGLAVLDFLLQMSAFVALLSLSKQRAS 805

Db 734 SMLFSFSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIKRQKN 793

Qy 806 RLDVCCCKVQBLPPPGQ--GEGLLIGFFQKAYAPFLHWTIRGVVLLFLALFGVSLYSM 864

Db 794 RLDIFCCVGAEDGTSVQASSELFRPFKNSYSPILLKDWMPRIPIAIFVGVLSFIAVL 853

Qy 865 CHISVGLDQELALPKDSYLLDYFLFNRYFEVGA PVYFVTTLGYNFSSBAGNNAICSSAG 924

Db 854 NKVDIGLQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEGRHDYTSKQGNMVCWGNG 913

Qy	925	CNNFSFTQKIQYATEFFEQSYLAIPASSWVDDFIDWLTP-SSCRLVYISGPNKDKPCPST	983
Db	914	CNNDLSVQOIFNAQAQLDNYTRIGFAPSSWIDDYDFWVKPQSSCCRV---DNITDQFCNAS	970
Qy	984	VNSLNCNKNCSIT-MGSVRPSVQFHYLPWFNDRENPNKCPKGLAAYSTVN--LTS	1040
Db	971	VVDPAACVR-CRPLTEGKQRPQGGDMRFLPMFLSDNPNPKCGKGHAAYSSAVNILLGH	1029
Qy	1041	DGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVGTDPAPAFVFPYITNV	1100
Db	1030	GTRVGATYFMTYHTVLOTSAFDALKKARLIASNT-ETMGINGS--AYRVFPYSVYV	1086
Qy	1101	FYEQYLTIIDDTIFNLGSLGAIPLVTVLLGCELWSAVIMCATIAMVNMFGVMWLMG	1146
Db	1087	FYEQYLTIIDDTIFNLGSLGAIPLVTVLLGCELWSAVIMCATIAMVNMFGVMWLMG	1146
Qy	1161	ISYNAVSLNLSVAGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVAGVAMTN	1220
Db	1147	ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFGITLTK	1206
Qy	1221	LPGLVLGLAKAQLQIIFFRMLNLLITLLGLHGLVFLPVILSYVGPDPVNA	1272
Db	1207	FGGIIVLAFAKSQIFQIFFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVNKA	1258

RESULT 24
US-10-736-731-4
; Sequence 4, Application US/10208731
; Publication No. US20030092038A1
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894

Qy	925	CNNFSFTQKIQYATEFFEQSYLAIPASSWVDDFIDWLTP-SSCRLVYISGPNKDKPCPST	983
Db	914	CNNDLSVQOIFNAQAQLDNYTRIGFAPSSWIDDYDFWVKPQSSCCRV---DNITDQFCNAS	970
Qy	984	VNSLNCNKNCSIT-MGSVRPSVQFHYLPWFNDRENPNKCPKGLAAYSTVN--LTS	1040
Db	971	VVDPAACVR-CRPLTEGKQRPQGGDMRFLPMFLSDNPNPKCGKGHAAYSSAVNILLGH	1029
Qy	1041	DGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVGTDPAPAFVFPYITNV	1100
Db	1030	GTRVGATYFMTYHTVLOTSAFDALKKARLIASNT-ETMGINGS--AYRVFPYSVYV	1086
Qy	1101	FYEQYLTIIDDTIFNLGSLGAIPLVTVLLGCELWSAVIMCATIAMVNMFGVMWLMG	1146
Db	1087	FYEQYLTIIDDTIFNLGSLGAIPLVTVLLGCELWSAVIMCATIAMVNMFGVMWLMG	1146
Qy	1161	ISYNAVSLNLSVAGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVAGVAMTN	1220
Db	1147	ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFGITLTK	1206
Qy	1221	LPGLVLGLAKAQLQIIFFRMLNLLITLLGLHGLVFLPVILSYVGPDPVNA	1272
Db	1207	FGGIIVLAFAKSQIFQIFFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVNKA	1258

Query Match 34.8%; Score 2402.5; DB 5; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.2e-199;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
7 RGLLWALLRLAQSEPTTIHQGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
4 RGLALGILLLLCPAQVFSQ-----SCVWYGEGLIAGDKRYNCEYSG----- 46
60 NTPARKITGPHLILLQKICPRLYTGNPTQACSAKQLVLEASISITKALLTRCPACSDN 119
47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDLQLPLQSLRCPSCFYN 103
120 FVNLHCHTCSNQSLFINTR---VAQLGAGQLPAVVAEYAFYQHSFAEQSYDSCSRV 175
104 LLNLFCBLTCSPROSQFLNVTATEDYDVPVNTQTKTNKELQYVVGOSFANAMYNACRDV 163
176 RVPAAATLAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDIT-----FHLLEPGQA 228
164 EAPSSNDKALGLLCKDADA-CNATNMLEYMFKNDNGOAPPTITPVPSPDPVH----- 215
229 VSGIQLPNEGVARCNBSQGGDVATCSQDCAASCAPAIARPO-----ALDSTFVLG 279
216 ---GWEPNNAATKGCDBSDEVTAPCSQDCQDSIVCGPKPQPPPPAPWTILGLDAMVIM 272
280 QMPGSLVLIILCSVFAVVTILL-----VGFRVAPARDKSMVDPKKGTSL 326
273 WITYMAFLVFPFGAFFAVWCYRKRYFVSEYTPIDSNATFSV-NASDRGE-----ASCC 324

US-10-756-149-4924
; Sequence 4924, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: File
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4924
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4924

```
; CURRENT APPLICATION NUMBER: US/10/208,731
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1319
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-208-731-4

Query Match      34.5%   Score 2385; DB 4; Length 1319;
Best Local Similarity 40.0%   Pred. No. 7.9e-198;
Matches 524; Conservative 231; Mismatches 440; Indels 116; Gaps 28;

QY 14 LLLRLAQSEBYTHIQGYCAPYDECGKNPELGSMLTSLNSVCLSNTPARKITGDHLIL 73
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 53 LLLLLCPAQVESQ-----SCWYGECC---IATGD---KRYNCKYSGPPKPLPKDGYL 100

QY 74 LQKICRPLYTGPNTQACCSAKOLVLSLEASITKALLTRCPACDFPNVNLHCHNTCSPNQ 133
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 101 VQSLCPGLFP--DNVSLCCDDIQQQLTKLSNLQLQLFLSRCPFCFYNLMTLFCBLTCSPHQ 159

QY 134 SLFINTRVAQLGAGOLPA---VVAEAPYQHSFAEQSYDSCSRVRVPAATAVLGWTMC 189
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 160 SQFLNATATEDYDPKTPENKTWKELEYVYGSGFANAMTACRDVEAPSSNEKALGLLC 219

QY 190 GYVGSALCNAQRWLNFGDGTNGLAPLDI-----TFHLLPQGAVGSGIQLNEGVARCN 244
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 220 GRDARA--CNATNWIEMFNKNDGQAPPTIIPVFSLSIL-----GMEPMENATKGN 270

QY 245 ESQGDVATCSCDCACPAIARPAQALDSTFYLGOMPSGLVLIILICSVFVAVTILLVG 304
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 271 ESDEVTVGSCQDCQDCIVCG--PKPQILQIGGGGW--GLDAMYVMWTVYVAFLEVFVFG 327

QY 305 FRVA-----PARDKSKMVDPKGTSLSDKLSFSFTHLLGQF 340
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 328 ALLAVNCHRRRYFVSEYTPIDSNIAFSVNSDDIGE-----ASCCDPLGAAPDDCLRRM 380

QY 341 FQCGTWVASWPLTILVSPVVALAAGLVFTLTPVELMSAPNSQARSEKAFHDOH 400
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 381 FTKWGAFCVRNPTCIIFSLAFITVCSGLVFQVTTNPVELMSAPHSQARLEKEFYDKH 440

QY 401 FGPFTNQVILTAPNRSSRYRDSLLGPK--NPSGILDLDLLELLELQERLHLQVWSP 459
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 441 FGPFFRTEQLIIQAPNTSVHIEPYEPAGADVPPGPPLNKEILHOVLNLQ-----IAIESI 495

QY 460 EAQRN---ISLQDICVAPLNPTSYLDCCINSLLQYFQNNRTLLLTANQTLMGQTSQV 516
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 496 TASYNNETVTDICVAPLSYPNK---NCTIMSVLNYFQNSHAVL-----DSQV 541

QY 517 -----DWKHFLYCANAPITFKDGTALALSCMADYGAPVPPFLAIGYKGYSEAE 568
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 542 GDDFYIADYHTHFLYCVRAPASLNDTSLHGPCLGTGFGVFPVPLVGLGYDDQNNAT 601

QY 569 ALIMPTSLNYPAGDPRLAQAKLEWAPLEEMRAFQORMAGMFQVFTFAERLSEDEINT 628
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 602 ALVITPPVNNYNDTERLQRAWAKEFTISFVKVYKN---PNLTISFTAESRIDEELNRE 658

QY 629 TAEDLPFATSYIVIFLYISLALGSYSSSRVWVDSKATILGGVAVVLGAVMAAGFPS 688
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 659 SNSDVFVTIISVYVWFLYISLALGHIOQSCSRLLVDSKISLGIAGLIVLSSVACSLGIFS 718

QY 689 YLGIRSLVILQVPPVPLSVGADNIFIVLEYQRLPRPGRPREVHIGRALGRVAPSNL 748
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 719 YMGMPLTLLVIEVIFPLVAVGVNDNIFILVQYQYQDERLQEBTLDOQLGRLIGEVAPTMF 778

QY 749 LCSLSEACFFLGTMPMPAVRTFALTSGLAVILDFLLQMSAFVALLSLDSKQREASRLD 808
```

```
Db 779 LSSFSETSAFFPGALSSMPAVHTFSLFAGMAVLIDFLLOITCFVSLIAGLDIKROEKHLD 838
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 809 VCCCVKPOELPPPOG---EGLLGFQKAYAPFELLHMITRQVWVLLLELFLFCVSLYSM 864
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 839 ILCCVRGAD---DQOQSHASESYLFRFFKNYFAPLLKOWLRPIVAVFVGVLSFSVAVV 895
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAIVYFVTTLGYNFSSEAGMAIACSSAG 924
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 896 NKVDIGLDQSLSPNDSYVIANFKSLAQYLSHSGPPVYFVLEBEGYNYSSRKQGNWVCGMG 955
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 925 CKNPSTQKIQYATEPEQSYLAIPASSWDDDFIDMLTP--SSCRLYISGNKDKFCPST 983
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 956 CNDLSLVQOIFNAAELEDTYTRVGFAPSSWIDDFWVSPQSSCCRLY---NVTHQFCVAS 1012
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 984 VNSLNLCKNCMSIT--MGSVRPSVEQPHKVLPELNDPRNPKCPKGGGLAAAYSTSVLTSDG 1042
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1013 VMDPTCVR--CRPUTPGKORPOQKBFMKFLFMSLNDPNPKCGKGGHAAVGSANVIGDD 1071
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1043 QVL--ASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPVITNVF 1101
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1072 TYIGATYFMTYHTILKTSADYTDAMKARLIASNITETWRS--KGSD--YRVFPYSVPYV 1128
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1102 YEOYLTPGLPMLSLCLVPTPAVSCLLLGLDLRLSGLMLLISVMILVDTVGFPMALWDI 1161
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1129 YEOYLTIIDTIFNLVSLSGIFLVTLVGLCWLMSAVIMCITIAMILVNMFGVWMLWGI 1188
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1162 SYNANSLINLAVSGMSVEFVSHITRSPALSTKPTWLERAKEATISNGSAVPAVAMTNL 1221
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1189 SLNAVSLVNLVMSCGISVEFCSHITRAFTMSTRSVSRABEALAHMGSSVFSGITLTKF 1248
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1222 PGLVTLGAKAQILQIFFFFRLNLLITLLGLLHGLVFLPVLISYVGPDPVNEA 1272
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1249 GGVLVLAFAKSQIFBIFYFRWYLAUWLLGATGLIFLPLVLLSYIGFSVAKA 1299
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 25
US-11-097-143-12003
; Sequence 12003, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-12003
```

Query Match	32.7%;	Score	2262;	DB	6;	Length	1287;
Best Local Similarity	37.8%;	Pred. No.	4e-187;				
Matches	504;	Conservative	244;	Mismatches	469;	Indels	116;
Gaps	30;						
Qy	10	LLWALLRLAQSEPTTHIQPCYCAFYDECGNPELSGLMTLSNVSCLSNTPARKITGD	69				
Db	17	LTAALVFLTIQSSKOD-----CWVYGVCNTN-DFSHS-----QNCYPNGTAKEMATD	62				
Qy	70	HLILQKICPRLYTGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDNDFNVLHCHNTC	129				
Db	63	GLELLKRCGFLLENSENKFCDDKNQVELLNKNVELAGNILDRCPCSMENLVRHICQPTC	122				
Qy	130	SPNQSLFINVTRVAQIGAGQLPAVVAYEAFYQHSFABOSYSCSRVRVPAATLAVGTWC	189				
Db	123	SPKQAEFMHVATQNKXGD-EYISSVDLHISTEYINKTYKSCQSVSPQTGQLADFLMC	181				
Qy	190	GUYGALCNAQWLFNCOGDTGNGLAPLDTTFHLLPQGAAGSGIOPLNEGVARCNEQGD	249				
Db	182	GAYSASRCNPTKWFNMGDATTNYPVFQITYIQHEP-KSNSNNFTPLNVTTPCNOAVSS	240				
Qy	250	DVATSCODCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSVPFV-VT	299				
Db	241	KLPACSCDCDLSGQGPPEPRPPEPKIVGLDAYF-----VIMAAVFLVGL	288				
Qy	300	ILLVG---PRVAPARDKSNVDPKKGTSLSDKLSFS-----THLLGQFFQGW	344				
Db	289	VFLMGSLFTQSSMDNDNQVD---GNDVSDEMPYSENDSYPEKLGHAHTETLETFTFTKW	345				
Qy	345	GTWVASMPLTILVSVIPVVALAGLVTETLTPVELWSAPNSOARSKAPHDQHPGF	404				
Db	346	GTYPASNPGTLTAGASLVILGYGINFTEITDPVKLWASNSKRLEREFDTKFPFP	405				
Qy	405	FRNQVILTAPNRSSVRYDSLLGPKNFSGIILDLLLELLELQERLRLQWSPSAQRN	464				
Db	406	YRLEQIIKAVNLPOIVHNT-SNGPYTFGPVDFREPLTKVLDLQEGIKEINA-----NG	458				
Qy	465	ISLQDICYAPLNDPNTSL--YDCINSLLOYFONNTLILLTANQTLMGQTSQVDWKDHF	522				
Db	459	TQLKDICYAPLSDGSEIDSVQVQSIWGYGDDRL-----DDHDEONGFNVTYLDAL	514				
Qy	523	LYCANAPLTFKDTALALSCMADYGAPVPFLAIGY-----KGKDYSEAEALIMTF	574				
Db	515	YDCISNPLY-----CLADYGGFVDPAPALGGLPFPDQGLTGSTKFPELANALITF	564				
Qy	575	SLANNYPAGDPRLAQAKLWEAFLEEMRAF-QRMAGMFQVFTFAERSLEDEINRTAEDL	633				
Db	565	LVKQH-HNKTDLENALTWEKKFVEFTNTYTKNMSQYMDIAFTSERSIEDELNRESQSDV	623				
Qy	634	PIFATSYIVIFLYISIALGYSYSSWSRVMVDSKATLGLGVAVVLGAVMAAMGPFYILGIR	693				
Db	624	LTILVSYLIMFMYIAISLGHKFKRFVDSKITLGTIGGVIIIVLASVSVSGVFGYIGLP	683				
Qy	694	SSLVILQVPELVLSVGDADNIFIVLEYORLPRRGPPEPRVHIGRALGRVAPSMLLCSLS	753				
Db	684	ATLIIIVEIPFLVAVGVNDNIFILVQHQDRQKPNETLEQQVQGRILGKVPKSLTSL	743				
Qy	754	EATCFEFGALTPMPAVRTFALTSGVLIDFLQMSAFVALLSIDSKRQASRLDVCCCV	813				
Db	744	ESFCFFLGGISDMFPAVAFALYAGVALLIDFLQITCFVSLFTLDTKREENRMDICCFI	803				
Qy	814	KPOELPPPGOGEGLLGFFQKAYAPFLHWHITRGVLLLLFLALPGVSLYSNMCHISVLGDQ	873				
Db	804	KGKKPDSITSNEGLLYKFFSVVPFLMKKIVRASVMVIFPAMLCFPSIAIAPRIDGLDQ	863				
Qy	874	ELALPKDSYLLDYFLNRPFEVCAPVYFTTLGYNESSEAGNNAICSSAGCKNFSPTQK	933				
Db	864	ELAMPQDSFVLHYFQSLNENLNIGPPYFVLKGLDAYTNSDDQLVNCAGQVCHNDDSVLTQ	923				
Qy	934	IQVATEPPEOSYLAI PASSWDDFDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNCL	990				
Db	924	IYLASHNSQTYIARPASSWIDDDYFDWAAASSCKY-----RKDSGDFCPH--QDTSCL	976				
Qy	991	KNCMSITMGS--VRPSVSEQHFKYLPWFLNDRNRIKCPKGGLAAYSTSVNLTSDDGQVL---A	1046				

Db	977	R--CNITKNSLLRBEKEFKYLPFLKDNPDPTCAKAGHAAYGAVKYSNSHERLNTEA	1034				
Qy	1047	SRFMAHKPKNSQDYTEALRAARELANITADLR-----KVP-GTDPAPAEVPPYITN	1099				
Db	1035	SYFMAYHTLKSSADYFLALESARKISANITQMLQGRILMSGVPMASALTVEVFPYSVY	1094				
Qy	1100	VFYEOYLTPLEGFPMLSCLVPTFAVSCILLGLDRLSLNLLSIVMLVDTVCGFMALW	1159				
Db	1095	VFYEOYLTMMSDITLQSMGIVLSIFPVTFVLMGDFVHSALVVVITIMVNLGSLMYW	1154				
Qy	1160	DISYNAVSLINLVAAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVMT	1219				
Db	1155	NISLNAVSLVNLWAVGIVSEFCHLVHSFATSKSVSQIDRAADSLSKMGSSIFSGITLT	1214				
Qy	1220	NLPGLVGLAKAQIQLIFFPRMLNLLITLGLLHGLVFLPVLSYVGVDPVNPALAEQKR	1279				
Db	1215	KFAGILVLAFAKSIQFQVFRMYLGIWVIGAAHGLIFLPVLLSYIGAPVSNARLRYHSQ	1274				
Qy	1280	A----EEAARAAM	1288				
Db	1275	AAAHETALAGIL	1287				
RESULT 26							
US-11-097-143-2679							
; Sequence 2679, Application US/11097143							
; Publication No. US20050208558A1							
; GENERAL INFORMATION:							
; APPLICANT: Venter, J. Craig							
; APPLICANT: et al.							
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID							
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE							
; FILE REFERENCE: DROSOPHILA GENES.							
; CURRENT APPLICATION NUMBER: US/11/097,143							
; CURRENT FILING DATE: 2005-04-04							
; PRIOR APPLICATION NUMBER: 60/157,832							
; PRIOR FILING DATE: 1999-10-05							
; PRIOR APPLICATION NUMBER: 60/160,191							
; PRIOR FILING DATE: 1999-10-19							
; PRIOR APPLICATION NUMBER: 60/161,932							
; PRIOR FILING DATE: 1999-10-28							
; PRIOR APPLICATION NUMBER: 60/164,769							
; PRIOR FILING DATE: 1999-11-12							
; PRIOR APPLICATION NUMBER: 60/175,693							
; PRIOR FILING DATE: 2000-01-12							
; PRIOR APPLICATION NUMBER: 60/184,831							
; PRIOR FILING DATE: 2000-02-24							
; PRIOR APPLICATION NUMBER: 60/191,637							
; PRIOR FILING DATE: 2000-03-23							
; NUMBER OF SEQ ID NOS: 43008							
; SOFTWARE: FastSeq for Windows Version 4.0							
; SEQ ID NO 2679							
; LENGTH: 1223							
; TYPE: PRT							
; ORGANISM: DROSOPHILA							
US-11-097-143-2679							
Query Match							
Best Local Similarity							
Matches							
27.0%; Score 1864.5; DB 6; Length 1223;							
33.5%; Pred. No. 1.7e-152;							
Conservative 237; Mismatches 498; Indels 113; Gaps 27;							
Qy	62	PARKTG-DHLLILQKICPRLY-----TGPN--TQACCSAKQLVSLSEASISITKALLTRC	113				
Db	14	PARPLNSPTSEAFKRCPLMYKEYKGESGEDELSCDDAAQIETWESGLSQADGVFSRC	73				
Qy	114	PACSDPNVNLHCHNCTSPNOSLFINVTRVAQIGAGQLPAVVAYEAFYQHSFABOS	169				
Db	74	PTCTRNWALTVCAMTCAKNHTLFLTAYNDTN-----DAGVDYVKYIDYRLTDDTVSKIY	127				

170 DSCSRVRPAAATLAVGTCMGVYGSALCNAQRWLNFGDGTNGGLAPLDTITFHLLEPGQAV 229
128 NSCIGIQHTGTGRPMDLGGCSYNAKTCNRRYRFPMDGSDYGVFPQINXKWSDEAE- 186
230 GSGIQPLNEGVARNCSQGDVATCSQDCAACPAIARPAQALDSTFYLGQMPG--SLVL 287
187 GSNEIYLDLSPKGGSEYDSA-CACIDEECSPLTDAFTGPDWLKWIAGLIGVTFILA 245
288 IIILCSVFAVVTTLLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQF----FOG 343
246 LIACALSFFIFWGAFGKTSAP-----SVCMTPLFGBFFYHGFRI 285
344 WGTWVASWPLTILVLSVIPVWALAGLVPTLTDVPVELWASPNQARSEKAFHDOHFGP 403
286 WGFPCAKHPVIVLALCSWATAGLSFGIRYNTITDPVELWAGESOTRIEKDYFDQHFGP 345
404 PFTNQVILTAPNRSSRYRSDLLGPKNFSGIIDLDDLLLELLELQERLRLHLOWMSPEAQ 463
346 FYRTNQMFVKAQVNTYFTHET-SNGVLNFGPAFEYNFLKEVFLQDSIMKLG-AD 399
464 NISLQIDICVAP-INPONTSLYD--CCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDH 521
400 NEGLDKICVAPVLMAGETPTVDRCALQSVYGVYFQHDMDRF--ENSVDNSNNYTYNLQ 456
522 FLYCANAPLTFKDGKTALASCMADYCAPVPPFLAIGY-----KGDYSEBALIWTFS 575
457 LEDCLRPM-MED-----CFGTGPIEPGIAVGMPKVAVGEDPDYMLATGLVLTEL 508
576 LNNYPAGDPLRAQAKLWERAFLERAFQRMAGMFQVTFPABRSLEDEINRTAEDLPI 635
509 GRNY-NDESKLEPMKWEKLVDFLDYK---SDRLDIATMAERSIQDAIVLSEGEVST 564
636 FATSIVIVILYISLALGYSXSSWRVMVDKATILGLGGAUVVLGAVMAAMGFYSYLGRSS 695
565 WISYVVMFYVAIALGHIRSCRGFLRESRIMLAIGGIVIVLASVCSLGFWGLDYVTT 624
696 LVILQVVPFLVLSVGADNIPFVLEYQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSA 755
625 MIAIEVIFPLVAVGVDNIFIMVHTYQRLDHSKFKTTHEAIGSAIQVGPSLIQTAGSEM 684
756 ICFPLGALTPMPAVRTFALTSLGLAVILDLQMSAFVALLSDLSKQESRLDVCCKVK- 814
685 ACFAICISDMPAVKTFAMTAAILDLFLQITAFVAMLAIDKRYLDGRDLMLCCKVS 744
815 -----PQELPPPGQGGELLLGFQKAYAPFLHWTIRGVVLLFLFALFGV 859
745 GKGKINDEGDGVDREKVE-----GLLETLPXNFYSPFLSKPKVSVLLIFTITCL 797
860 SLYSMCHIISVGLDQELALPKDSVLLDYFLFLNRYFVGAPVYFVTTLYGNFSEAGMNAI 919
798 SLWVTPSIEKGLDQEMSPKNSHVVKYFRYVMDLLANGAPVYVWLKPLNYSLEPLQONLI 857
920 CSSAGCNNSFTQKIQYATEPQSVLAIAPASSWDDFDLWTPSSCCHLYISGPNKDXF 979
858 CGGVECNNSLSVQLYTOAQYBEITSARPPASSWLDYIDLWLAISDCCRYNT---TGGF 914
980 CFSVTNSLNLCKNCSITMGSVRPSVEQFKYLPWFLNDRPNLCKPGGLAAYSTSVNLT 1039
915 CSNSKSEDLCPCERGFTENGRLPDAETFNKIPIYFLPDLPAEACKAGASTADAVIYT 974
1040 SD----GOVLASRFMAYHKLKNSQDYTBALARAELAANITADLRKVFGTDPAPEVFPY 1095
975 IDDVGMSTVQDSYFMQYSTSTTSBEEFYSQLREVRIISGEINAMPKE---NNVDAEIFAY 1031
1096 TITNPFYEQYLILPGLPMLSLCLVPTFAVSCLLLGLDLRGLLMLLSIVMLVDVTFQF 1155
1032 CVFYIYEQYLTWGDAMFSLGMSLVAIFLVLTLLITGLDITSTFVLFMVICILINMLGM 1091
1156 MALWDSYNVLSLNLVAVGMSRFPVSHITRSEFISTKPTWLBRAKEATISMGSAVPAG 1215
1092 MWAISINLNAISLVNLVUCVGVGEFVAHIVRSFK-RABGTQAERARHSLNTGSSVLGS 1150
1216 VAMTNLPGLVLGLAKAQLIQIFFFLNLILITLLGLLHGLVFLPVLISLVGPDVNPALAL 1275

1151 ITLTKFAGIVVLGFSNSQIFQVYFRMYLIGVILGAAGLILLPVLLSLGPPQK-----L 1206
1276 EOKRAEEAVALVAVAS 1291
1207 ARSSGABPTASITITT 1222

RESULT 27

US-10-208-731-6
; Sequence 6, Application US/10208731
; Publication No. US20030092038A1
; GENERAL INFORMATION:
; APPLICANT: Carstea, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/10/208,731
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-208-731-6

Query Match 19.2%; Score 1329; DB 4; Length 1170;

Best Local Similarity 28.8%; Pred. No. 8.1e-106;

Matches 373; Conservative 230; Mismatches 512; Indels 182; Gaps 46;

10 LLWAL-----LLRLAQSEPTTTHQGYCAFYDECGKNP-----ELSGSLMTLSNVSCLSN 60
3 VLWIIALVQLARLVQGV-----TATCAMYGNCCKSVFQGNELPCVPSPSEPPVLS 54
61 TPARKITGDHLLIQLKICPLRYTGNPTQACCAKQLVLSIASLITKALLTRCPACSDNF 120
55 ETSK-----LLVEVCGEWEKVR-YACCTKQVVALRDNLQKAQLISSCPACLKNF 105
121 VNLCHNTCSNOSLFINVTRVAQLGAGALPAVVAEAFVQHSFASQSDSCSRVVPAA 180
106 NNLFCFTCRADQGRFVNITKV-EKSKEDKDIVAELDFWNSSWASEFTDSCNIIKPSAT 164
181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGN--GLAPLDTITFHLLEPGQAVSGIQPLNE 238
165 N----GYAMDLLIGGAKNYSQLFKLGDAKPMLGSPQINIKYDLANEE--KEWQEFND 218
239 GVARNESGDDVATCSCODCAACPAIARPAQALDSTFYLGQMPG---SLVLIILCSVP 295
219 EYVACDDAQ---YKACSDCQSCPHL-KP-LKDGVCVKVPLPCFSLSVLIFFYITCALF 272
296 AVVTILLVGFVRVAPARDKSMVDPKKGTSLSDK-----LSFSTHT-----LLQRFQCGW 345
273 AFMYTYLCKRKKNGAMIVDDDIVPSSG-SLDESETNVFSPFNNETNFPNGKLANLTKVYG 331
346 TWVASWPLTILVLSVIPVWALAGLV-FTELTTDPVELWASPNQARSEKAFHDOHFGP 404
332 QFSVENPKLITTVFSIFVPSFIIFQYATLETDPILNLWVSKNSEKFEKEVFDNFGFP 391
405 FRTNQVILTAPNRSSRYRSDSL--LLGPKNFSGIIDLDDLLLELLELQERLRLHLOWMSPEAQ 462
392 YRTEQIFVFNNETGPNVLSYETLHWWFDVENF-----ITEEL-----QSS 429
463 RNISLQIDICVAPLNDPNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDH 522
430 ENIGYQDLICFRP-TEDST----CVTESFTQYFQG-----ALPNKDSWKREL 470
523 LYCANAPLTFKDGKTALASCMADYCAPVPPFLAIGYKGYKGYSEBALIWTFSNLNPAG 582

Db	471	QBCGKFP-----VNCUPTFOQPKTNLL-----FSDDDLNNAHAFVVTLLLTNH--- 514	;	ORGANISM: Caenorhabditis elegans	US-10-208-731-9
Qy	503	DPRLAQAKLWEBAFLEENRAFORMMAGMFQVFTTAERSLEDEINRTTAEDLPFATSIV 642	Query Match	15.6%; Score 1078; DB 4; Length 1296;	
Db	515	-----TOSANRWEER-LEEVLLDLKVEGL-RISFNTFISLERELNN--NDISTVAISYM 567	Best Local Similarity	24.0%; Pred. No. 7.3e-84;	
Qy	643	IFLYISLALGSYSRSMVMSKATLGGGVAIVGAVMAAGPFYSYIGIRSSLVLQVW 702	Matches	321; Conservative 237; Mismatches 492; Indels 288; Gaps 41;	
Db	568	MELYATWALRRKDGKTRLL-----LGISGLLIVLASIVCAAGFLTFLGLKSTLIAEVI 621			
Qy	703	PELVLSVGADNIFIVLEQVR-LPRRPGPREVHIGRALGRVAPSMWLLCSLSEACFFLG 761			
Db	622	PFLIIAIGDNIIFLITHEYDRNCEQKEPYSIDOKIISAIGRMSPSILSLLCQTCFLIA 681			
Qy	762	ALTPMPAVRTFALTSGLAVIDFLQMSAFVALLSDSKRQASRLDVCCCVKQPELPPP 821			
Db	682	AFVTPAVNFAIYSTVSIFNGVLQLTAYVLSILSYEKRSYKQIT----- 728			
Qy	822	GQEGELLGFFQKAYAPFLLHWITRGVILLFLALFGVSLYSNMCHISVGLDQELALPKDS 881			
Db	729	-GNEETKESFLKTFYFKMLTQ--KRLIIIFSAMFPTSFLVFLPEIQFGLDQTLAVPQDS 784			
Qy	882	YLLDYFLNRPFEVCAPYFVYTLGYNFSSEAGMNAICSS-AGCNNESTFKIQIYATEF 940			
Db	785	YLVDFYKDVYSFLNVPVYVYVVK-NLDTKRNQOQKICGKFTTCERDSLAVNLE---OE 840			
Qy	941	PEOSYLAIPASSWVDDFIDWLTSS--CCRLYISGPNKDKFCPSTVNSLNCNKMCSITM 998			
Db	841	RHRSTTEPLANWLDIYFNLNDQNDQCEL-----KKGTEVCPSPSPSRRC-ETCFQ--Q 894			
Qy	999	GSVR-----PSVEQPHKYLPMFLNDRPNKCPKGLAAYSTSVNITSDGOVLASRFMAY 1052			
Db	895	GSNNYNSMGPECKDFMEYLSIWIN-APSDCPLGGRAPYSTAL-VYNETSASVFRTA 952			
Qy	1053	HKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPVTITNVFQVYLTILPEG 1112			
Db	953	HHPLRSQKQFIQAY-----SDGVRISSFFPELDMFAYSPPYIFVQIUGPLT 1001			
Qy	1113	LFMLSLCLVPTFAVSCILLGLRLSGLNLLSIVMILVDTVGPMAWDISYNAVSINLV 1172			
Db	1002	LKLGSAIILIFISSVFL-QNIRSSFLLALVTMIIVDIALMALLGISNAVSLVNL 1060			
Qy	1173	SAVGMSVEFVSHTRSPAI---STKPTWLERAKEATISMGSAVFAVGMNTNLPGLVLGL 1229			
Db	1061	ICVGLGVEFCVHIVRSFTVVPSETKKDANSRVLSYSLNTIGESVIKGITLTKFTIGVCVLA 1120			
Qy	1230	AKAQLIOIFPRNLNLLITLGLLHGLVFLPVILSYVG 1266			
Db	1121	AQSKIFDVYFRMWFTLIIIVAAHLHALLFLPALJSLFG 1157			
RESULT 28					
US-10-208-731-9					
; Sequence 9, Application US/10208731					
; Publication No. US20030092038A1					
; GENERAL INFORMATION:					
; APPLICANT: Caratea, et al.					
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease					
; FILE REFERENCE: 4239-53894					
; CURRENT APPLICATION NUMBER: US/10/208,731					
; PRIOR FILING DATE: 2002-07-29					
; PRIOR APPLICATION NUMBER: US/09/462,136					
; PRIOR FILING DATE: 2000-06-01					
; PRIOR APPLICATION NUMBER: PCT/US98/13862					
; PRIOR FILING DATE: 1998-07-02					
; PRIOR APPLICATION NUMBER: US 60/051,682					
; PRIOR FILING DATE: 1997-07-03					
; NUMBER OF SEQ ID NOS: 13					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 9					
; LENGTH: 1296					
; TYPE: PRT					

```
Db 879 VYHDPN--TFCSTNRKNSALDKACKTCMDP----- 908
QY 1025 PKGLAAYSTSVNLTSQGVASRFMAYHKPL--KNSQDYTEALRAARELANITADLRK 1082
Db 909 --DGRASFKAISRTSGRIQASQFWTFHKKLSISNSSDFIKAMDYARWYSRRLERSI-- 964
QY 1083 VPGTDPAFVPPYTIITNVFYEQYLITLPEGLFMLSCLVPTPAVSCLLGLDLRSGLNL 1142
Db 965 ----DPTARHFAVSKFPFPFYEQYSTIMPILTQFTVVGFGVICVTLGIDVKGACAV 1020
QY 1143 LSIVMILVDTVGPMLWDISYNVSLNLVSAGMSVEFVSHITRSFAISTKPTMLERAK 1202
Db 1021 ICQVS-----NYFHVSS--GILIEFVNVLKGACSLRQAKDRAE 1060
QY 1203 EATISGSAVPAAGVAMTNLGIILVLGLAKAQLIQIPFFRLNLITLLGLHGVFLPVIL 1262
Db 1061 STVSGIPILSGPVVMTAGSTWFLSGAHLQIITVYFFKLFLITIVSSAVHALIILPILL 1120
QY 1263 SYVGPDVNPALAEQKRAEAAVAVMASCNPNSRVSSTADNIYV-----HSFEGS 1314
Db 1121 AFGSRGHGSGSETSTNDNDQHDACVLS--PTAESHSINVEEGILNRPSSLLDASHILDPL 1178
QY 1315 IKGAGALSFLPNNGROF 1332
Db 1179 LKAEGGIDKAI----GRDF 1193

RESULT 29
US-10-239-316-8
; Sequence 8, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: KITAYAMA, Yoshio
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same And
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 8
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-8

Query Match 13.6%; Score 942; DB 4; Length 194;
Best Local Similarity 99.5%; Pred. No. 2.7e-73;
Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 615 FTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGYSWSRVMVDSKATLGGVA 674
Db 1 FNAERSLEDEINRTAEDLPFATSYIVIFLYISLALGYSWSRVMVDSKATLGGVA 60
QY 675 VVLGAVMAAGPFSYIGIRSSVILQVLPVLVSVCAADNIFIVLEYQRLPRPGPREV 734
Db 61 VVLGAVMAAGPFSYIGIRSSVILQVLPVLVSVGADNIFIVLEYQRLPRPGPREV 120
QY 735 HIGRALGRVAPSMMLCSLSAICFFLGCALTMPAVRTFALTSGLAVIDFLQMSAFVAL 794
Db 121 HIGRALGRVAPSMMLCSLSAICFFLGCALTMPAVRTFALTSGLAVIDFLQMSAFVAL 180
QY 795 LSLDSKQKQASRLD 808
Db 181 LSLDSKQKQASRLD 194

RESULT 30
US-10-424-599-211862
```

```
; Sequence 211862, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211862
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
US-10-424-599-211862

Query Match 13.0%; Score 900.5; DB 4; Length 492;
Best Local Similarity 39.3%; Pred. No. 4.9e-69;
Matches 199; Conservative 99; Mismatches 170; Indels 39; Gaps 11;

QY 736 IGRALGRVAPSMMLCSLSAICFFLGCALTMPAVRTFALTSGLAVIDFLQMSAFVAL 795
Db 1 ISIALVEVGPSITLASVSEVLAPAVGFSFISMAIRVFSMFAALAVLLDFLQVTFVALI 60
QY 796 SLDSKQKQASRLDVCVCPQBLPPPGQE---GLLGPFQKAYAPFLHWTIRGVVLL 852
Db 61 VLDSLRADKRVDCPCIKVHADPDIGTGRKRGPLLARYMKVHAPILSITGWKIVIAI 120
QY 853 FLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYEVEGAPVYFVTTLGNFSS 912
Db 121 FVGFALASALTSTRIEPLGLEQIVLPDSDYLOGYFNNVSEYLRIGPPVYFVVK-NTNYSS 179
QY 913 EA-GMNAICSSAGCANNFSTQIQVATEPPEQSYLAIPASSWDDDFIDWLTPSS--CCRL 969
Db 180 ESTHTNQLCSISHCNSDSLNEIARAALVPDTSYIAKPAASWLDLDFLVWSPFAFGCCRK 239
QY 970 YISGPNKDKFCP-----STVNSLCLKNCMSI-----TMGSVRPSVEQFHKYLP 1013
Db 240 FTNG-----SYCPDDQPPCCAPGSSCVSGTKDCTCFRHSDLHNDPSTTQPREKLP 295
QY 1014 WFLNDRPNIKCPKGLAAYSTSVNLT--SDGQVLASRFMAYHKPLKNSODYTEALRAARE 1071
Db 296 WFLSSLPSADCAKGGHGAYTSSVELKGYDNGIIQASSFRYHTPLNKQVDYVNSMRAARE 355
QY 1072 LAANITADLRKVGTDPAPFVPPYTIITNVFYEQYLITLPEGLFMLSCLVPTPAVSCLL 1131
Db 356 FSRVSDSLK-----IEIPYSVYFMYFPFQYLIHWKTALINLAIGAIVIV-CLIF 406
QY 1132 GLDLSGLNLISIVMILVDTVGFPMALMDISYNAVSLINLVSAGMSVEFVSHITRSFAI 1191
Db 407 TSSLSWSSSIILLVLAIVVDLMGLMAILNLQNLALSVNLVMSVGIAGEFVCFVHMTSFTV 466
QY 1192 STKPTWLERAKETISMGSAVFAGVAM 1218
Db 467 ASGDR-DORAKEALGTMGASVFSGITL 492

RESULT 31
US-10-424-599-189288
; Sequence 189288, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211862
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
US-10-424-599-211862
```



```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 189288
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_141944C.1.pep
US-10-424-599-189288

Query Match          12.4%; Score 859.5; DB 4; Length 541;
Best Local Similarity 35.3%; Pred. No. 2.1e-65;
Matches 197; Conservative 109; Mismatches 187; Indels 65; Gaps 14;

QY 752 LGEATCFGLGATPMPAVKTEALTSLGVLDFLLQMSAFVALLSLD--SKRQEA SRLDV 809
Db 2 LSCRVC-----SMIPAPAFRIY-----YIPLITAFVALITFDIFISAYMESFASFP 46
QY 810 CCCVKFQELPPPCQGE-----GLLGFFQKAYAPFLHMTITRGVVLLFLALFGVS 860
Db 47 CVMILNP---PSAERNEGTCVRRERDGLITRYMKVEHAPFLGLCGVKILVIAVFAAFTLAS 103
QY 861 LYSMCHISVGLDELALPKDSVLLDYFLFLNRYFEVGPVYFVTTLGYNFSSEA-GMNAI 919
Db 104 IALCTRIEPLGELQALPDRSLYQGFYSNISEYLRVGPPLYFVVK-DYNSLESKHTNQL 162
QY 920 CSGAGCNFSFTQIOYATEFPEQSYLAIPASSWDDFDLWLTSPS--SCRLYISGPNKD 977
Db 163 CSISHCDNSNLNEISRAISLVPTSSVIAPASWLDLDFLWISPEAFSCCRKF---TWD 218
QY 978 KFCPTVNSLNCILKNCMSTMGSV-----RPSVEQFHKYLPWFNLDRPN 1021
Db 219 SYCPDDQPPCLPDRGPGCLGGCKDCTTCFRHSLDVLNDRPSTAQFREKLEFWLDALPS 278
QY 1022 IKCPKGLAAYSTSNLT--SDQVLA SFMAVHKELKNSQDYTEALRAARELAANITAD 1079
Db 279 ADCAKGHHGAYTNSVLDNGYEGVIOQASEFRYHTTPNQGDKVNAIRARDFSAKISS 338
QY 1080 LRKVPGTDAPEFPPFTITNVFYEQYLTILPEGLFMLSCLPPTFAVSCLLGLDLRLSG 1139
Db 339 LK-----MDIFPYSVFIFEQYLDIWKALINISALGAIFV-CLITSSVWSSV 389
QY 1140 LNLISVIMLVDTVGFMAWMDISYNAVSLINVSAGMSVEFVSHITRSPAISTKKTWLE 1199
Db 390 IILLVIMILDLGMVMAILGIQINAVSVVNLMSIGIAVEFCVHIVHAFTVSLGDR-SQ 448
QY 1200 RAKENTISGSAVPAGVAMTNLPGLVGLAKAQLIQIPFRNLTLTLGLHGLVFLP 1259
Db 449 RAKTALCTMGASFSGITLTKVGVVLCTSTSQIFVYIFQMYLALVLGLHGLVFLP 508
QY 1260 VILSYVGPDPVNPALALEQ 1277
Db 509 VVLSLFGPPLRYTVIKEQ 526

RESULT 32
US-08-954-701A-19
; Sequence 19, Application US/08954701A
; Publication No. US20030032085A1
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

```
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-701A-19

Query Match          9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.4e-46;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQAVGSGIQPLNEGVARCNESQGDVATVCSQDCAASCPAIARQALDSTPYLQMPQS 284
Db 24 FGRPAGGRRRTGGLRAAAPDRD-----YLHRPSYCDAAFALEQI--- 65
QY 285 LVLIILICSVFAVVTILLVGRFVAPARDKSKWDPK-----KTSLSDKLSFSHTLL 337
Db 66 LVLIILICSVFAVVTILLVGRFVAPARDKSKWDPK-----SKGATGRKAPLWRAKFORLLFKIGCTYIQKNC 98
QY 338 GQFFQMGWGTWASWPLTILVLSVIPWALAAGLVFTETTDPELWSAPNSQARSEKAPH 397
Db 99 GKF-----LVGGLLIFGAPVGLKKAANETNVEELVVEGGRVSRRLNT 143
QY 398 DQHPGPPFRFN-QVILTAPNRSRYDSLLGLPKNFSGLDLDLLELE--LQERLRHL 454
Db 144 RQKIGEEAMFNQLMIQTPKBEG-----ANVLTEALLQHLDSALQASRVHV 190
QY 455 QWVSPEAQRNISLQDICYP-----LNPDTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 191 YMYN----RQWKLEHLCKYKSGELITETGYMDQIIEYLYPCLITPLDCFWEGAKLSGTA 246
QY 506 NQTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 ---YLLGKPLRWNTNPDPLEFSELKKINQVDSWEEMLANKAEVGHGYMDRPLNADPDC 304
QY 528 -----APL-----TFKQGT-----ALALSCHMA 544
Db 305 PATAPNKNSTKELDMALVNLGCHGLSRKYMHWQEBELIVGGTVKNSTGLKLSAHLQTMF 364
QY 545 DYGAVPFPFLAIGGYKDYSEAEALIMTFPSLNNYPAGDPRLQAQAALWEAEFLEEMRAFQ 604
Db 365 QLMTPKQMY---EHPKGYEY-----VSHINNE-----DKAAALAEAWQRTYVEVHVQS 411
QY 605 RRMAGMFQVFTTAERSLEDEINRTTAEDLPFIFATSYIVIFLYTSLALGSYSSRSRVNVD 664
Db 412 AQNSTQKVLSTFT--TTLDILKSFSDSVIRVASGYLLMLAYACLTW---LRWD--CSKS 465
QY 665 KATLGLGGVAVLVGAMAAAGPFSYLGIRSSVLQVPPFLVLSVGADNIFIFVLEQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNATTTQVLPFALGVGVDDVFLLAHAFSET 525
QY 725 PRPGEPRVHIGRALGRVAPSMGLCSLSAICFFLCAITPMPAVRTFALTSLGLAVLDF 784
Db 526 GQNKRIFFEDRTGECLKRTGASVALTISNVTAFMAALIPALRAFSLQAAVVVVVNF 585
```


Db	914	INPSA--FYIYLTAWNSNDPVAYAAQANIRPHRPWVHDK---ADWMPETRLRIPAAEP	968
Qy	1044	VLASRFMAYHKLKNSODYTEALRAARELANITA-DLRKVPGTDPAFVPPYITITNVFY	1102
Db	969	IYEAQFPYINGLDRDTSDFEAEIKVTRICNSYISLGLSSYPNG-----YPP-----LFW	1018
Qy	1103	BOYLITLPEGFMISLCIVPTFAVSCLLGLDLSGLLNLSIVMILVDTVGFMAIWDIS	1162
Db	1019	BOYIGLRHLLFISVVLACTFLVCAVELLNPTAGII-VMVLAIMTVELFGMGLIGIK	1077
Qy	1163	YNVSLINLVSAGMSVEFVSHITRSP--ALSTKPTWLERAKEATISMGSAVEAGVAMTN	1220
Db	1078	LSAVPFWIIASVGIVGEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST	1133
Qy	1221	LPGLIVLGLAKAQLIQIEFFRLNLILTLGLLHGLVFLPVLISYVG--PDVNPALALEQ-	1277
Db	1134	LLGVLMLAGSPDFIVRFVFAVLAITLTLGVNLGLVLLPVLSPFGPYPEVSPANGLNRL	1193
Qy	1278	--KRAEEAAVAVMVASCPNHPHSRVSTADNIYVNHSPFGSIKG	1317
Db	1194	PTPSPPEPPSVVRFAMPFGHTH--SGSDSDSEYSSQTTVSG	1233
RESULT 34			
US-09-754-032-19			
; Sequence 19, Application US/09754032			
; Publication No. US20030148388A1			
; GENERAL INFORMATION:			
; APPLICANT: SCOTT, MATHEW P			
; GOODRICH, LISA V			
; JOHNSON, RONALD L			
; TITLE OF INVENTION: Patched Genes and their Use			
; NUMBER OF SEQUENCES: 19			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert			
; STREET: Four Embarcadero Center, Suite 3400			
; CITY: San Francisco			
; STATE: CA			
; COUNTRY: US			
; ZIP: 94111			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/754,032			
; FILING DATE: 03-Jan-2001			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/540,406			
; FILING DATE: 06-OCT-1995			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Rowland, Bertram I			
; REGISTRATION NUMBER: 20015			
; REFERENCE/DOCKET NUMBER: a60190-1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415-781-1989			
; TELEFAX: 415-398-3249			
; INFORMATION FOR SEQ ID NO: 19:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1447 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:			
US-09-754-032-19			
Query Match 9.4%; Score 651.5; DB 3; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.4e-46;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			

Db 969 IEVAQPFYINGLRDTSDFEALIEKVKRTICSNYTSGLSSYPNG-----YFP-----LFW 1018
QY 1103 EQLYTLPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSIYMIIVDTVGFMAWDIS 1162
Db 1019 EQYIGLRHLLFISVVLACTFLVCAVFLNPTAGII-VWVLALMTVELFGMMGLIGK 1077
QY 1163 YNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLERAKGATISMGSAVAGVAMTN 1220
Db 1078 LSAVPVVILIASVIGIEFTVHVALAFLTAIGDKN---RRAVLALHEMFAPVLDG-AVST 1133
QY 1221 LRGILVLGAKAQLIQIFFRRLNLLITLGLLHGLVPLPVLSVVG--PDVNPALALEQ- 1277
Db 1134 LLGVMLAGSEFFIVRYFFAVAILITLGLVNLGLVLLPVLSPFFGYPPEVSPANGLNRL 1193
QY 1278 --KRAEEAANVMVASCPNHPSPVSTADNIYVNHSPSGSIK 1317
Db 1194 PTPSPPPPSVVRPAMPGTH--SGSDSDSEYSSQTTVSG 1233

RESULT 35

US-10-421-446-19

; Sequence 19, Application US/10421446

; Publication No. US20030186309A1

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GOODRICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fiehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/421,446

; FILING DATE: 22-Apr-2003

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/724,631

; FILING DATE: 28-No. US20030186309A1-2000

; APPLICATION NUMBER: US/08/656,055

; FILING DATE: 1996-05-31

; FILING DATE: <Unknown>

; FILING DATE: 08/540,406

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1447 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-10-421-446-19

Query Match

Best Local Similarity 9.4%; Score 651.5; DB 4; Length 1447;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGAVSGIQPLNEGVARCNESQDDVATCSCODCAASCPAIARPOALDSTFYLGOMPGS 284
Db 24 PGRPAGGRRRTTGGRLRAAADRD-----YLHRPSYCDAAFALEQI--- 65
QY 285 LVLIILILCSVFVAVVTILLVGFVAPARDKSKMVDPK-----KGTSLSKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWRAKFORLLFLKGCYIKNC 98
QY 338 GQPFQCGMTWASWPLTILVLSVIPVVALAAGLVFTTETDPVELMSAPNSQARSEKAFH 397
Db 99 GKF-----LVVGLLIIFGAFVAGLKAANLETNVEBELVVEVGGVRSRELYT 143
QY 398 DQHFPGFPFRTN-QVILITAPNRSRYDLSLLGPKNFGIILDLLELLE--LQERLHL 454
Db 144 RQKIGBEAMFNPQLMIQPKBEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QWSPQAQRNLSLODICYAP-----LNPNTSLYDCCINSLLIYQFONNRTLLLLTA 505
Db 191 YMYN-----RQWKLHLCYKSGELITETGYMDQIIEYLYPCLITITPLDCFWEGAKLQSGTA 246
QY 506 NOTLMGO-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFNFDPLFLELKKINYOVSWEENLKAKEVGHGYMDRPLNPADPDC 304
QY 528 -----APL-----TPKCGT-----ALALSCMA 544
Db 305 PATAPNKNSTKPLDMALVNLGGCHGLSRKYMHWQBELIVGGTVKNSTGKLVSAHALQTMP 364
QY 545 DYGAVPVFPFLAIGGYKGYDSEBALIMTFSLNNYPAGDPRLAQAKLWEAPFLERAFQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE---DKAAAILAEAWORTYVEVHVQSV 411
QY 605 REMAGMFQVTFATERSLEDEINRTAEDLPIFATSYIVIFLYISLALGSVSSSRVWVDS 664
Db 412 AQNSTQKVLSTFT--ITLDDILKSFSDSVIRVASGYLLMLAVACLTW---LRWD--CSKS 465
QY 665 KATLGLGGVAVVGLVAMAAAGFFSYLGRSSVILQVVPFLVLSVGADNIFIVLEYSQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFALGVGVDDVFLLAHAFSE 525
QY 725 PRPGEPREHVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPMPAVTPTALTGLAVILDF 784
Db 526 GQMKRIPFEDRTGECLEKRTGASVALTSISNVTAFMAALIPIPALRAFSLQAAVVVVVF 585
QY 785 LLOMSAFVALLSLDSKQOEASLDVCCC-----VKQEL-----PPG 822
Db 586 AMVLLIFPAILSMGLYRREDRLDIFCCFTSPCVSRVIOVBFQAYTDTHTNTRYSPPPY 645
QY 823 QGEGGL----- 827
Db 646 SSHSFAHETQITMQSTVQLRTEYDPTHVYVYTABPRSEISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGFFOKAYAPFLHWHITRGVLLLLFLALFGVSL 861
Db 706 SSTDRLLSQFSDSLHLCLEPPCTKWTLSFAEKHYAPFLKKPKAKVWVIFLPLGLIGVSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA PVYFTTILGYNFSSEAGNNAICS 921
Db 766 YGTTVRDGLDLDIVPRETREYDPIAAQFKYFSP-----YNN----- 803
QY 922 SAGCNNSFTQKIQYAT-----BPPEQSYLAIAPASS-----WVDDFDIDLW----- 961
Db 804 -----YIVTQKADYFNIOHLLYDLHRFSFNKYVMLEENKQLPKWMLHYFRDWLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDKPCPSTVNSLNCNKCMSTMG 1000
Db 858 AFDSDWETGKIMPNNVKNGSDGVLAYKLLVQTSRDK--PIDISOLT--KQRLVDADGI 913
QY 1001 VRPSVEQPHKYL-PWFLND-----RNIKCPKGLAAY--STSNLTSDQ 1043
Db 914 INPSA--FYIYLTAWVSNPDVAYAAASQANIRPHREWHDK---ADYMPETRLRIPAASP 968

FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 60/019,765
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Laurence J.
REGISTRATION NUMBER: 35, 551
REFERENCE/DOCKET NUMBER: 015280-278200US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1296 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..1296
OTHER INFORMATION: /note= "amino acids encoded by human
nevoid basal cell carcinoma syndrome
(NBCCS) (PATCHED (PTC)) cDNA"
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-10-302-279-60

Query Match 9.1%; Score 628; DB 4; Length 1296;
Best Local Similarity 23.6%; Pred. No. 1.3e-44;
Matches 254; Conservative 165; Mismatches 384; Indels 272; Gaps 39;

QY 426 LLG--PKNFSGLDLDLLELLEQLERLRHLQVMSPEAQRNIS-----LDICYPALNP 477
DB 97 LLGKPLRWTFNDPFLFEELKKI-----NYQVDSWEEMLNKAESVGHYMDRPCLNADP 151
QY 478 DNTSLYDCCTNSLLQYQNRTLLLLLTANTQTLMGTSQ-VDMKHFLYCANAPITFKDGT 536
DB 152 D-----CPATAPKNKSTKPLDMALVLNGCGHLSRKMYMQUEBILVGG----TVKNST 200
QY 537 -----ALALSCWADYGAVPFPELAIGGYKGDKDYSEAEALIMTFSLLNNYPAGDPRLAQAKL 591
DB 201 GKLVSAHALQTMFLMTPKQMY---EHPFGVEY-----VSHINWNE----DKAAAILAEA 247
QY 592 WEEAFLEMRAFORRMAGMFQVTTFAERSLEDEINRTTAEDLPFPATSYIVIFLYISIAL 651
DB 248 WQRTVVEVHVQSAQNS'QKVLSTFT--TTDLILKSPSDSVIRVASGYLLMLAYACLTM 306
QY 652 GSYGSWSRMVDSKATLG LGVVAVLVGA VMAMGFYSYLGRSSLIIVLQVPFLVLSVCA 711
DB 307 ---LRWD--CSKSQAGVLAGVLLVALSVAAGLGLCSLIGISFNAAATTQVLFFLALGVGV 361
QY 712 DNIFIPVLEYQLPRPGEPREVHIHGRALGRVAPSMLLCISEALCFFLGALTMPPAVRT 771
DB 362 DDVFLLAHAFSGTGQNKRIPIFDRTGECKRTGASVALTISNVTAFFMAALIPALRA 421
QY 772 FALTSGLAVIDFLQMGAFVALLSIDSKROEASRLDVCCC-----VKPQEL-- 818
DB 422 FSLQAUVVVFNFAMVLLIFPALLSMDLYRRDRRLDFCCCTSPCVSRVIQVEFOAYTD 481
QY 819 -----PPPGQGEGL----- 827
DB 482 THDNTRYSPPPPYSSHSPAHEQTITMQSTVQLRTEYDPTHVYVYTAAEPSRSEISQPVTV 541
QY 828 -----LLGFFOKAVAPFLLHWITRGV 848
DB 542 TQDTLSCQSPESTSSTRDLLSQFSDSSLHCLBPCTKTWTLSSFAEKHYAPFLLKPXKV 601
QY 849 VILLFLALFGVLSYMCHTISVGLDQBELALPKDSYLLDYFLNLNRFEVGPVYFVTTLGY 908
DB 602 VIPLFLGLGLGVSLYGTTRVRDGLDLTDIVPRETREYDFIAAQKFYSF-----Y 650
QY 909 NPSSEAGMNAICSSAGCNNFSFTQIKQYAT-----EPPEOSYLAIPASS-----W 953
DB 651 NW-----VIVTKADYPNIOHLLYDLHRSFNSNVKYVMLENKQLPKMW 693


```
QY 560 -----KGKDYSE-----ARALIMTFSLNNYP-----AGDPLAQAKLWEEAPLEE 599
DQ 182 GSSLGWVTKMGENETDILASAKAWFMIVHLKFPBEMSYISGE-----WE---LEL 230
QY 600 MRAFORMAGMP-QVFTFAERSLEDEINRTTAEDLPIEATSYIVIFLYISLGSY--SS 656
DQ 231 GRMLTQVPEDPYISITIFHSQTLADELKRNADTLIPRIISITLLIVFSLCSLSPIDS 290
QY 657 WSRVMVDSKATLGLGGVAVVGLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSVGDNIFI 716
DQ 291 FSDWLSKFLSILGWSAGIALLTGVGFLSLMGMPYN-DIVGVMPFLVAVGVDMFEL 349
QY 717 FVLEYORLPRPCEPREVH--IGRALGRVAPSMMLCSLSEAIFFIIGALTPMPAVRFPAL 774
DQ 350 MVAAV-----RRTSRTHVHERMGCELAADAASILTSTDVLSFGVGAITTIPAVQIFCV 405
QY 775 TSGLAVIDFLQMSAFVALLSLDSKQEAR-----LDVCCCVK-----814
DQ 406 YTGVAIFFAFYIQTTPFAACLAMKHEASGRNSLFLIEAVSABKTSISTFORLNLGS 465
QY 815 -POELPPGOGGELLGFFOKAYAPFLHLHWITRGVLLLF-LALFGVLSYSMCHISVGLD 872
DQ 466 VPDHSASHDVKPLTSRPFGEWAPVLMHPVRGIAWVFIYLLGAS-YGCSRIKEGJE 524
QY 873 QELALPKDSYLLDYFLFLNRYF-EVGAPVYFV-----TTLGYNFSSSE 913
DQ 525 PWNLLVEDSYAIPHYRLLEKFWKYGOQVQVINNAPDLRNHTSRDRVHAMVLDPATSKH 584
QY 914 A-GMNAICSSAGCNFPSTQKQYATEPPEQ-----SYLAIPASSWDDFDIDL 961
DQ 585 AIGMESV-----QFWLFEMERYQKELEVOIIDSFYGLLHHFLASKTNNPLAEDIYW- 637
QY 962 TPSSCCRLYISGNKDKPCFSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFLNDREN 1021
DQ 638 -----GMPD-----642
QY 1022 IKCPKGGLAAYSTSVNLTSBGQVLAS-RFMAHYKPLKNSQDYTEALRAARELANITADL 1080
DQ 643 -----DDNGTMVKSFRFILGMDLVTTMDQTDATMSPREVAARW----681
QY 1081 RKVPGTDPAFEPFPYITIVFVEQYLTILPEGLFMLSCLVPTFAVCSLLGLDLRSGLL 1140
DQ 682 -----PEFNVTFMPIWMFTDQYIIITENTVQNIILALLVMIVIAVLFIQPMCS-LW 733
QY 1141 NLLSIWMLVDVTFGMALWDISVNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLER 1200
DQ 734 VALACASIDFGVIGYMTLGMVNDLAIMTITIMSGISFVDYSAHAYGVVVSREDTAAGR 793
QY 1201 AKEATISMGSAVFAGVAMTNLPGLVILGLAKAQLIQIFFRMLNLITLLGLHLGVFLPV 1260
DQ 794 VKEALSALCWPLSQG-AMSTIAVSVLADI PAYMI-VTFKTVVLSISGLHLGVFLPV 851
QY 1261 ILSYV-----GPDVNPALALBOQRABEA-----VAAMVASCNPNRVS 1300
DQ 852 LLSIFVRGCCIIIPSSPHGPSAQKIEKQIRIAAISSSPLDLRTVAPLRASSPISPHRLE 911
QY 1301 -TADNLYNHSPEGSTK 1316
DQ 912 YDESPTVHNRSKNSIK 928
```

RESULT 42

```
US-10-890-776A-4805
; Sequence 4805, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
```

```
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 4805
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Consensus sequence of Patched motif
US-10-890-776A-4805

Query Match      8.6%; Score 597; DB 5; Length 821;
Best Local Similarity 23.7%; Pred. No. 3.2e-42;
Matches 223; Conservative 169; Mismatches 368; Indels 180; Gaps 35;

QY 366 LAAGLVFTLTTPDVPWELMSPNSQARSEKAFHQHDFG-----PFRNTQVILTAPNRSYR 421
DQ 2 LSKRYT-FTDDIRSYTERSGRSEHEPLVERKPPGQGYPIRFS--IITAKAN---55
QY 422 YDSLLGPKNFGSGLDLDLLELLELQERLRLQVNSPEAQRNISLQDICYAPLNDPTS 481
DQ 56 -----SNILDIRYNEVQVNDPI-STNVITPVSGITLSPKDL------93
QY 482 LYDCCINSLLQVFNRRITLLLTANQTLMGQTSQVDWKDHFLYCANAPITFDKGTALALS 541
DQ 94 MRFCINCPVNOHYNGQ-----ILRSNL-----HNSRID-----LTYP 126
QY 542 CNADYGAPVPPPLAIGGVK-----GKDYSEABALIMTFSLNNYPAGDPLAQ--AKLWEEA 595
DQ 127 TMTFPTKIYLPNGFGVKLDENPGNIKSVKAIIVLYLKR-----DPEVEDSKREWLS 182
QY 596 FLEEMRAFQRRMAGMFQVTTAERSLSEDEINRTTAEDLPIEATSYIVIFLY-----ISLA 650
DQ 183 LFDYLE--NEYASDHQVTFPSDQVLEDELVRNGLTLTPFFVVGPAITVPSILTSVRLA 240
QY 651 LGS-YSSWSRVWDSKATLGLGGVAVVGLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSV 709
DQ 241 SGLSHIDWVR-----SKPILAILGLVLTPLMAIVSAFGLLFWLGRFNISIVC--VMPFLVLAI 295
QY 710 GADNIFIFVLEYORLPRPCEPREVHIGRALGRVAPSMMLCSLSEAIFFIIGALTPMPAV 769
DQ 296 GVDNFMFLMHAWQRTTRSLSVKK--RMGEVLEAGPSITITSLTNVLSFGIAGITPTPEI 353
QY 770 RTFALTSGLAVIDFLQMSAFVALLSLDSKQEAR-----LDVCCCVKQELPPPG 822
DQ 354 QLFCTYTAIVAFDFIYQITFPAAIMTAGYKMEGRHSLFLRAVDAEETSPQRDLSRG 413
QY 823 -----QGEGLLGLFPKQAYAPFLHLHWITRGVLLLF-LALFGVLSYSMCHISVGLDGE 874
DQ 414 SAKSKSHEAQVAVVKKFLLNIYCPFLNPKVRVCVLLVYVLAIAIYCTNNKKGLOPA 473
QY 875 LALPKDSYLLDYFLFLNRY-EVGAPVYFVTTLGYNFSSSEAGMNAICSSAGCNNSFTQK 933
DQ 474 KLFKDSPLVEYLRUREKHVMPYGLQVTVFNNPPDLINPENRDL-----519
QY 934 IQYATEPPEQSYLAIPASSWDDFDIDLMTTPSSCCRLYISGNPKDKPCFSTVNSLCLKNK 993
```

Db	520	-----NMWDEPEN--TP-----YAMGKNSTKFW-----LRDYENP 548
Qy	994	MSITGVRPSVEQ-FHKYLPWFNDRENIKCPKGG-----AAYSTSVNLTSDGQVLAS 1047
Db	549	LYSPFISELEDEEWFYDLLEWFLKS-PGFSHWGDLWMDNKTDYEITI-----VKCF 600
Qy	1048	RFMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVFPYTIINVFYEQVLT 1107
Db	601	RFTTGGKDLSTWDRTRLLKTKRWGADBY-----PDFNVTVDEDAFLDQILS 649
Qy	1108	ILP-----EGFLMSLCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAWDISYN 1164
Db	650	IGPTTQSIWTWLCM-----AVVCELPINPNTVFVITVSTASIDIGVGFSLMGVDLD 705
Qy	1165	AVSLINLVSAVGSVEFVSHITRSPAIS--TKPTWLERAKEATISMGSAVFAGVAMTNLPG 1223
Db	706	PISMITIISIGFSDVSAHIAHYFVRSHGSETPDERLADALEALGWPIQ-AALSTILC 764
Qy	1224	ILVLGLAKAQLIQIFFRNLNLLITLLGLHLGLVFLPVILS 1263
Db	765	VLPLLFVPSYMV-VVFFKTFIFLVVWVIGLHGLIFLPIILS 803
RESULT 43		
US-10-060-756A-3		
; Sequence 3, Application US/10060756A		
; Publication No. US20030046717A1		
; GENERAL INFORMATION:		
; APPLICANT: Zhang, Jian		
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN		
; FILE REFERENCE: PB0177		
; CURRENT APPLICATION NUMBER: US/10/060,756A		
; CURRENT FILING DATE: 2002-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00667		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00664		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00669		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00665		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00668		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00663		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: US 09/864,761		
; PRIOR FILING DATE: 2001-05-23		
; PRIOR APPLICATION NUMBER: US 60/327,898		
; PRIOR FILING DATE: 2001-10-09		
; NUMBER OF SEQ ID NOS: 4804		
; SOFTWARE: Aecomica Sequence Listing Engine		
; SEQ ID NO 3		
; LENGTH: 954		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-10-060-756A-3		
Query Match 8.5%; Score 588; DB 4; Length 954;		
Best Local Similarity 21.8%; Pred. No. 2,4e-41;		
Matches 219; Conservative 190; Mismatches 356; Indels 240; Gaps 39;		
Qy	334	HT-----LLGQFPQGWGTVASWPLTILVSVIPVVALAAGLVF--TELTTDPVELWSAP 386
Db	115	HTDCLEGLSRTFQWLQWQGAHPWIFLLAPLMTAALGTGFLYLPKDEEDLEEHTPV 174
Qy	387	NSQARSEKAFHQHGFPPFTNQVLLTAPNRSS-----YRYSLLLGKPNFSGIL 436
Db	175	GSPAKAERRFVQGH-----FTTNDSYRFSASRRSTEANFVLLVVSYSDSLDPAFVUS 230
Qy	437	DLDLLELLELLEQLERLHLQVMSPEAQNISLQDIC--YAPLNPDNTSLYDCCI--NSLLQ 492
Db	231	KLDGAVQDLRV-----AREKSGIQYQQVCARYAL-----CVPPNPILY 270
Qy	493	YFQNNRTLLLTANTQILMGQTSQVDMKHFYLCANAPLTFKDGTTALASCMADYGA-PVF 551
Db	271	AWQNKTLNL-----SSISPAY-----NHRGPLY 296
Qy	552	PFLAIGY-----KGQDYSEAEALIMTFSLNYPAGDPRL-AQAKLWEAEFLBEMRAP 603
Db	297	LTGFGGYYILGGSLMGQLLRKAMRLLYLK---TEDPEYDVQSKQWLTHLLDQFTNI 353
Qy	604	QRMA--GMQVVFYA-ERSLEDEINNTAEDLPFATSYIVIFLYISLALGSSYSSWSRV 660
Db	354	KNILALKKIEVWHFSLSRQLEPEATSVV--IPVFLHAYLILILFATVSCFRD----- 406
Qy	661	MVDSKATLGLGGVAVILGAVMAAMGFYSYLGRSLRSLVILQVVPFLVLSVAGDNIFIVLE 720
Db	407	CIRNMCAAFGVISAFNAVVGFGLLHIGV-FPVIIVANSFPLILGVGDDFMISA 465
Qy	721	YORLPFRPGEPREVIHGRALGRVAPSMLLCSLSAISAEFFLGALTPMPAVRTPALTSLAV 780
Db	466	WHK-TNLAGDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSVQCFYITGTTL 523
Qy	781	ILDFLQMSAFVALLSDSKQEASRLDVCCVKPQELP-----PPQGG 824
Db	524	LFCYFNITCFGAFMADGKEV-----VLCWLKADPKWPSKPKCCFPFGSVDEHG 578
Qy	825	EGL--LLGFFQKAYAPFLLHMITRGVLLLFALFGVSLYSKMSCHISVGLDQELALPKDSY 882
Db	579	TDIHPMSLFFRDYFGPFLTRSESKYFVFIIVLYIISIIYGFVHQEGLDLRLNLASDDSY 638
Qy	883	LLDVPFLNRYF-EVGAPVYFVTLLGYNFSSEAGMNAICSSAGCANNFPTQKIQYATEFP 941
Db	639	ITPYFNVEENYFSDYGPVWVIVTKVDYWDK-----DVRQKLENCIKIP 683
Qy	942	EQSYLAIP--ASSWVDDFDIDMLTSSCCRLVYISGPNKDKFCPSTVNSLCLNKMCSITWG 999
Db	684	EKNVYVDKNLTFEFLDAYVQ-----YLGNSQD---PNEKVT----- 717
Qy	1000	SVRPSVEQFHKYLPWFNDRENIKCPKGGLAAYSTSVNLTSDGQVLASRFMAYHKPLKNS 1059
Db	718	-----FMNNIPDLSNFPN-----FQHDINISSNEIISRGFTQTTDVSSS 759
Qy	1060	QDYTEALRAARELANITADLRKVPCTDPAFVFPYTIINVFYEQVLTLP----- 1111
Db	760	-----AKKI---LLFQLRRI-AEDCOIPLAVYNOAFIYFDQYAAILEDTVRNVLVA 807
Qy	1112	--GLFWLSLCLVP-----TFVSCLLGLDLRSGLNLLSIVMILVDTVGFMAWDI 1161
Db	808	SAAMFIVSLLIPLCLSLWVTFALGSVIGV-----TGFWAFKV 848
Qy	1162	SYNAVSLINLVSAVGSVEFVSHITRSPAISATKPTWLERAKEATISMGSAVFAGVAMTNL 1221
Db	849	NLDSISMINLVICTGFSFDSAHISYAFVSSQPSVNQKSVEALYLLGYFVLQS-AISTI 907
Qy	1222	PGILVGLAKAQLIQIFFRNLNLLITLLGLHLGLVFLPVILSVG 1266
Db	908	IGVCVLAARAKAVIFRT-PFKIMFLVMIFGAAGHLPIFIPVFLTFFG 951
RESULT 44		
US-10-890-776A-3		
; Sequence 3, Application US/10890776A		
; Publication No. US20050129683A1		
; GENERAL INFORMATION:		
; APPLICANT: Zhang, Jian		
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN		
; FILE REFERENCE: PB0177		
; CURRENT APPLICATION NUMBER: US/10/890,776A		
; CURRENT FILING DATE: 2004-07-14		
; PRIOR APPLICATION NUMBER: US 10/060,756		
; PRIOR FILING DATE: 2002-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00663		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00664		
; PRIOR FILING DATE: 2001-01-30		

Db	403	LLMLVACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGICALLGITFNAATQ 457
Qy	701	VPFPLVLSGADNI FIFVLEYQRLPRRPGEPREVHIGRALGVAPSMLLCSLSEAICPFL 760
Db	458	VLFPFLALGIVDDVFL--LAHAFTALPGTFQIQRMGECIQRGTGTVLSTINNNAFLM 515
Qy	761	GALTMPAVRTFALTSGLAVIDLFLQMAFAVALLSLDSKROEASRLDVCCC----- 812
Db	516	AALVPDPALRAFLQRAIVGCTFVAVMLVPAILSLDLRRHCORQLDVLCCFSFSPCSAQ 575
Qy	813	---VKPQEL-----PPPQOQEGL-- 827
Db	576	VIQILPQELGDGTVPGVIAHLTATVQAFTHCEASSQHVVITLPPQAHLPVPPSPDLGSEL 635
Qy	828	-----LLG-----FFQKAYAPFLAHMITRGVLLLLFLA 855
Db	636	FSPGGSTRDLLGQBEETROKAAKSLPCARWNLAHFARYQFAPLLQSHAKAIVLVIFGA 695
Qy	856	LFQVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAVPVFVTTGLYNFS-SEA 914
Db	696	LLGLSLYGATLVQDGLALTVDVPRGTEKHAFLSAQLRYFSL-YEVALVTCGGEDYAHSQ 754
Qy	915	GMNAICSSAGCWNFSFTQKIYATEPPEQSYLEIPAS-----SWDDFDIWLTPSSCC--- 967
Db	755	AL-----FDLHQRFSSL-----KAVLPPPTAQAPRTHLVYRNWLQIGIAAFDQ 798
Qy	968	-----RLYISGPNKDFCPSTVNSLCKNCHSITMGSRVPSVEQHKLPWFLND 1018
Db	799	DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQEPLD-FSQTTRKRLVD 846
Qy	1019	RPNIKCPKGLAAASTSVNLTSDQVLAASFMAVHKD-----1055
Db	847	REGLIPE--LFYVGLTVWVSSDFPLGAASQANFYPPPEHLHXDYTTGENLRIPPAQ 904
Qy	1056	-----LXNSODYTEALRAARELAANI--TADLRKVPGTDPAFEPVPTITNVFY 1102
Db	905	LEPAQFPFLRLGLQKTADFVEAIEGARAACAEGACQAGVHAYPSGPE-----LFW 954
Qy	1103	EQYLTILPEGLFMLSCL--LVPTFAVSCLLGLDLRLGLLNLLSIVMLIVDTVGFMAIWD 1160
Db	955	EQYLG--RRCFLLAVCTLLVCTFLVCALLLNPNWTAGLI-VLVIANMTVELFGIMGFLG 1011
Qy	1161	ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
Db	1012	IKLSAIPVILVASVIGVEFTFVVALGF-LTTGGSRLRAHALEHTFAPVDTG-AIST 1069
Qy	1221	LPGILVLGLAKAQLTQIFFFRLNLILITLLGLHLGLVFLPVILSYXGPDVNPALAEOKRA 1280
Db	1070	LLGILLMAGSHDFIVRYFFAALTVLITLLGLHLGLVLLPVLSILGPP--PEVIQMYKES 1127
Qy	1281	EEAVAAV-----MVASCPNHPSRVSTADNIYN 1308
Db	1128	PEILSPAPPOGGGLRWGASSLPUQSFARVTTSMTVAIH 1165

RESULT 46

[illegible]

```
Db 929 GARAACTEQAGVHAYPSGSPF-----LFWEQVGL--RRCFLLAVCIILLVCTP 976
Qy 1125 AVSCLLGLDLRSLNLLSIYVILVDTVGFMALMDISYNAVSLNLSYAVGMSVEFVSH 1184
Db 977 LVCALLLLSPWTAGLI-VLVLAMTVLFGINGFLGKLSAIPVVLVASIGVFEVTVH 1035
Qy 1185 ITRSPAISTKPTWLERAKATISMSANVAGV---AMTNLPGLVILVGLAKAQLIQIPPR 1241
Db 1036 VALGFLTSHSGSNLRAA-----SALEQTPAPVTDGAVSTLLGLLMLAGSNFDFIIRYFV 1090
Qy 1242 LNLTLTLLGLLGLVFLPVLSYVGP 1267
Db 1091 VLTVTLLGLLRLGRLLLPVLLSILGP 1116

RESULT 47
US-10-425-115-328081
; Sequence 328081, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 328081
; LENGTH: 343
; TYPE: PRN
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_62274C.1.pap
US-10-425-115-328081

Query Match 8.4%; Score 582.5; DB 4; Length 343;
Best Local Similarity 35.8%; Pred. No. 1.5e-41;
Matches 126; Conservative 71; Mismatches 116; Indels 39; Gaps 10;

Qy 848 VVLLFLALFGVLSYCMCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYFVTLG 907
Db 7 VVITVFGVLTASIALSTRQLGLEQIVLPNSVLDYDFDLAKYMKVGPPLYFVWK-D 65

Qy 908 YNFS- EAGMNAICSGAGNPFQTKIQYATEPPEQSVLAIPASSWDDFDNLTPTS- 965
Db 66 FNYASVDVTNQICISQCNSSLNLEISRQSLSPETSYIAKPAASWLDLFLNMSPEAF 125

Qy 966 -CCRLVIGPNKDKFCP-----STVNSLNLKNCMSITWGS---VRPSV 1005
Db 126 GCCKRFVNG-----SYCPDDPPCQQLDQVSGSCMTKTC-SNCTTCFLHSLDNGRST 180

Qy 1006 EQFKYLPWFLNDRPNIKCPKGLAAYSTSVNLT--SDGQVLASRFMAYHKPLKNSQDYT 1063
Db 181 TQFRDKLPWFLDALPSSDCSGKAGYSTSLDSGVSGIIOASAFRVTHTPLNKQSDYV 240

Qy 1064 EALRAARELANITADLRKVPCTDPAFVPPYITNVFQYLTILPEGLFMLSCLVPT 1123
Db 241 NSMARARDFSSKMSRDQ-----MKIFFYSVFYIPPEQYLSVMKTAIWNICVCLGTI 292

Qy 1124 PAVSCLLGLDLRSLNLLSIYVILVDTVGFMALMDISYNAVSLNLSYAV 1175
Db 293 FVV-CFIVTSSLWASAILVTVLAMVLMDMGVMAILGILQNLNATSVNLMVSI 343

RESULT 48
US-09-909-280A-2
; Sequence 2, Application US/0909280A
; Patent No. US20020160375A1
; GENERAL INFORMATION:
; APPLICANT: Buncroft, David A.
```

```
; TITLE OF INVENTION: HUMAN PATCHED GENES AND PROTEINS, AND USES RELATED
; TITLE OF INVENTION: THERETO
; FILE REFERENCE: CIBT-P02-050
; CURRENT APPLICATION NUMBER: US/09/909,280A
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 09/207,857
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 60/067,940
; PRIOR FILING DATE: 1997-12-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1203
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-909-280A-2

Query Match 8.4%; Score 580; DB 3; Length 1203;
Best Local Similarity 22.9%; Pred. No. 1.8e-40;
Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;

Qy 306 RVAPARDKSKMVDPKKGTSLSDKL--SFTHTLLGQFFOG-----MGTWVASWPLTILVLS 359
Db 3 RSPPLRELPSPSYTPPARTAAPQILAGSLKAPLWLRAYFOGLLFLSCGCIQRHCGKVLFLG 62

Qy 360 VIPVVALAAGLVFTBLTDPVELWGNAPNSQARSEKAFHDQHFQ--PFFRTNQVILTAPNR 417
Db 63 LLAFGALALGLRMAIETNLEQLWVEGSRVSRVQELHYTKELGERAAVYSQMLIQATQGE 122

Qy 418 SSRYDSLLLPKPNFSGILDLDLLELLELQRLHQLWQVSPQARNISLDICY---AP 474
Db 123 GEN-----ILTPEALG-----LHLQAALTASKVQVSLYKGSWDLNKICYKSGVP 166

Qy 475 LNPDN-----TSLVDCCINSILQVFNQNRITLLLTANQTLMGQTSQVDWMD- 520
Db 167 LIENGMIERIEKLPFCVILTPDLCFWGAK---LQGSAYLPGRPDQIQTWINDPEQLLE 223

Qy 521 -----HFLYCANAPLTPKDGTTALALSCMADYGA 548
Db 224 ELGPPASLEGPRELLDKAQVQVYGRPCLHPDDLHCPPAPNHHSRQAPNVAAHELSGGC 283

Qy 549 PVPPP-----LAIGYKGDYSE---AEALIMTP-----SLNNYPAGDPRLAAQKL 591
Db 284 HGFSHKPMHOBELLGGWARDPOGELLRAEALQSTFLMSPRQLYHEFRGDTYQTHDIGW 343

Qy 592 WEEAFLEEMRAFORERWAGMFVTFPABERSLE-DEINRTTAEDLPFPATS- 640
Db 344 SEEQASTVLAQWRFPVLAQALPENASQOIHAFTSTTLDI-LHAFSEVSAARVVG 402

Qy 641 IVIFLYISLALGSYSSSRVWVDSKATILGLGVAVVLGAVMAAGFFSVLGRSSILVIQ 700
Db 403 LLMLAYACVTM---LRWD--CAQSOGSVGLAGVLVALAVASGLGLCALLGITFNAAVQ 457

Qy 701 VVPLVILSVGADNIFIVLEYORLPRRPGEPREHIGRALGRVAPSMILCSLSEACPL 760
Db 458 VLPLGILGIGVDDVFL--LAHAFTALPGTLPQRMGECLQRTGTSTVLTSTNNMAAPLM 515

Qy 761 GALTMPAVRFPALTSGLAVIDPFLLOMSAFVALLSLDSKQESRLDVCCC----- 812
Db 516 AALVEIPALRAFSLQAAIIVGCTFVAVMLVPPAIFSLDLRRRHCRQLDLVLCFSSPSCAQ 575

Qy 813 ---VKPQEL-----PPKQGEGL-- 827
Db 576 VIQILLPQELGDGTVPVGLIAHLTATVQATHCEASSQHVVTILPQAHLVPPSPDPLGSL 635

Qy 828 -----LLG-----PFQKAYAPFLLHMTITRGVILLFLA 855
Db 636 FSPGSTRDLLQGEETROKAAKSLPCARWNLAHFARYQFAPLQLQSHAKAIVLVLFCA 695

Qy 856 LFGVSLYSWCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYFVTLGYNPS-SEA 914
Db 696 LGLSLYATLVDQGLALTVDVVRPGTKBHFASLAQLRYFSL-YEVALVTQGGFDVAHSOR 754
```


Db 5 DLEHYTPVGSAPKAERFPVQGH-----FTTNDSTYRFSASRRSTENAFVSLILVSVSDSL 60
Qy 428 GPKNFGSGLDLDLLELELOERLRLQWSPQAQRNLSLQDIC--YAPLPNDNTSLYDC 485
Db 61 DPATAEVSKLDGAVDLRV-----AREKSGSIQYQVCARYAL----- 100
Qy 486 CI--NSLIQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCNAPLTFKGTALALSCM 543
Db 101 CVPPNPILYAMQVNTLNL-----SSISFPAY----- 127
Qy 544 ADYGA-PVPPFLAIGY-----KGKDYSEAEALIMTFSLNMYNYPAGDPLR-AQAKLWEE 594
Db 128 -NHRHPLYLTFGFGYILGSLGMLQLLRKARMLLYLK--TEPEYDVQSKQMLT 183
Qy 595 AFLEEMRAPQRMA--GMQVFTTA-ERSLEINRTTAEDLPFATSYIVIFLYISAL 651
Db 184 HLLDQFTNLIKALKKIEVHTSLSRQLEFEATSVTV--IPVFHLAYILLPAVTS 241
Qy 652 GSYNSWRVMDSKATLGLGGVAVLGAAMAAGFFSYLGRSSLVILQVVPFLVLSGA 711
Db 242 PRFD-----CIRNKMCAAFGVISAFVAVSGFGLLHIGV-PFVIIANSFPLILGVG 295
Qy 712 DNIFFVLEYQRLPRPGPREVHIGRALGRVAPSMILCSLSEACFFLIGALTTPAVRT 771
Db 296 DDMFIMISAWHK--TNLAGDIRE-RMSNVYKAAVSTITITITNIALYTGIMSFRSQVC 353
Qy 772 FALTSGLAVIDLFLQMSAFVALLSLDSKROBASRLDVCCCVKQBELP----- 819
Db 354 FCITYGTLLFCYFYNITCFGAPMALDGRREV-----VCLWLKKADPKWPFKKFCPP 408
Qy 820 ----PQOQEGEL--LLGFQKAYAPFLHWTIRGVVLLFLFLALFGVLSYKMSCHISVLGD 873
Db 409 FGSVPDEHGTDIHPMSLFRDYGPELTSSEKSYFVYFVYLYVLISSYICGHVQEGDL 468
Qy 874 ELALPKDSYLLDYFLFLNYP--EVGAPVYFVTTLGVNFSSEAGMNAICSSAGCNPFPQ 932
Db 469 RNLASDYSYITPYFVNEENYFSDYGRVMVITVKVYWDK-----DVRQ 513
Qy 933 KIQYATEFPEQSYLAIP--ASSWVDYFIDWLTSPSCCRLYISGPNKDKFCPSTVNSLNC 990
Db 514 KLENTCKIPEKNVYVKNLTFFWLDAYVQ-----YLGNSQD----PNEKNT---- 556
Qy 991 KNCMSITMGSVRPSVQPHKYLFWFLNDRPNIKCPKGLAAYSTSVNLTSDGOVLASRF 1050
Db 557 -----FMNIPDFLSNFPN-----FQHDINISSNEIISRGF 589
Qy 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYTTITNVFYEQYLITLP 1110
Db 590 IQTTDVSS-----AKKI--LLFQLRRI-AEDCQIPLMVYNQAFIYFDQYAAILE 637
Qy 1111 E-----GLFMLSCLVP-----TPAVSCLLGLDLRSLGLNLLSVMLVDT 1152
Db 638 DTVRNVLVASAAMFIVSLLIPIPLCSLWTFPAIGSVIVGV----- 678
Qy 1153 VGFMAUWISYNAVSLINLVASGVMSVEFVSHITRSPATISKTWTLEAKBATISGSAV 1212
Db 679 TGFMAEFKNLSDISMINLVICTGSPDFSAHISYAFVSSSQBSVQKSYEALYLLGYVP 738
Qy 1213 FAGVAMTNLPGILVLGLAKAOLIQIPPRFLNLITLLGLLHGVFLPVLVILSVYG 1266
Db 739 LOS-AISTIIIGVCVLAANKAYIFRT-PFKIMFLVMIFGAAGHLIFIPVLTFFG 790

RESULT 52

US-10-890-776A-4808
; Sequence 4808, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A

; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acemica Sequence Listing Engine
; SEQ ID NO 4808
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: Sequence of the Patched motif of gi 1707052
US-10-890-776A-4808

Query Match 7.8%; Score 539.5; DB 5; Length 846;
Best Local Similarity 23.0%; Pred. No. 3.4e-37;
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;
Qy 401 PGPFRFTNQVIL-----TAPNRSSRYDSLLGPKNFSGILDLLELLE 446
Db 10 FGYPSTYERRIHDAPLVGDGTAVAGRAVTS--REVQVAVVARSGGNILDVRVFNELKL 67
Qy 447 LQERLHLQWSPQAQRNLSLQDICYAPLPNDNTSLYDCINSLLYQYFQNNRTLLLTAN 506
Db 68 MESFIRN-NITVQFNRTWSFADLCLIA--GPDGR---CANNDHIQ----- 106
Qy 507 QTLMGQTSQVDMKDHFLYCNAPLTFKGTALALSCMADYGAPVFPFLAIGGYK----- 560
Db 107 --LAKRLHQ-----HGINITPTVRLSDKSAYIAS-----ALGGVKLAKGN 146
Qy 561 GKD-YSEAEALIMTFSLNYP-----AGDPLAQAKLWEEAFLEEMRAPFORWAGMFOV 613
Db 147 GENIIVEATAMLLIYQLKFPYNEISYVSG-----LWREFKNQNDY-KKQAKYISI 197
Qy 614 TFTAERSLEINRTTAEDLPFATSYIVIFLY-----ISLALGSYSSWSRVMDSKAT 667
Db 198 TYFHSQTLSDENLRNAERLAPKFIAGFVILVCFSLCSIVTIKSGSYIDW---VVTKPI 253
Qy 668 LGLGGVAVVLGAAMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLEYQRLPR 727
Db 254 LSVLGVSNAGMGSIASAMGMLTYLEIQYN-DIIAVMPFLVAVGTDMFLMVASLKRTR- 311
Qy 728 PGEPREVHIGRALGRVAPSMILCSLSEACFFLIGALTTPMPAVRTFALTSGLAVIDLFLQ 787
Db 312 -NLKYDQRIAEACMADAASVILITALTALDSFGVGTITITPAVQIFCIYTWCAILLTFAYQ 370
Qy 788 MSFAVALLSLDSKROE-----ASRLDV---CCCVKPBQLPPPGQG--- 824
Db 371 LTFEFCALYYTTRIEBQGLHSIWLRPVATVSYSTSPNLVKLFWLGSQPK-PLPSCGVSS 429
Qy 825 -----EGLLGFQKAYAPFLH-WITRGVLLFLALFGVSLYS 863
Db 430 TSSVSTMTSQTSPASKHLHCAATSFRRNWPVLMQPMI-RAIAGLWYLIYLGISYIG 488
Qy 864 MCHISVGLDQELALPKDSYLLDYFLFLNR-YFEVGAPVYFVTT-----LGVNFSSE 913
Db 489 CTHLKEGLEPANLLVDSDSTATPHRYVLEKHYWHYGASLIQIVSVNPPDLRPPVERINWDM 548

QY 914 AGNNAICSSA-GCNFSP-----TQKQYATE-----PPEQSYLAIPASSWDDFI 958
Db 549 ASTFANCKVAIGDSSVQFWLREMQVSEIHKIQYDNEKFYDAAQYIYSDMSQPVWVDV 608
QY 959 DMLTPSSCCRLYISGPNKDKFCPSTVNSLCKNCKMSITMGSVRPSVEQFHKYLPWFLND 1018
Db 609 -----WGRNN 613
QY 1019 RPNIKCPKGLAAYSTVNLTSQGLASRFMAYHKPKNSQDYTEALRAARELANITA 1078
Db 614 -----NSERIILKTRFMIGMRDITTTKQTEATWTFREIASRF-- 651
QY 1079 DLKVPCTDPAPEVFPYITNVEYQYLILPEGLFMLSCLVPTTAVSCLLGLDLRG 1138
Db 652 -----EQNVITYMPLWLFDTQYALWVNTMODIIVAVACMLVISALLIPQVCSF 702
QY 1139 LLNLLSIVMLVDVTGFMALWDISYNAVSLINLVSAGMSVRFVSHITSPALSTPTWL 1198
Db 703 WV-AVTIGSIDGLVGLFMTLWNLDAISMITIIMSVGFSVDYSAHITVYISKESTTS 761
QY 1199 BRAKEATISMGSAAVAGVAMTNLPGILVLGLAKAQLIQIFPRLNLLITLLGLHGLVFL 1258
Db 762 ARVCDALGLWNPVAGQ-AMSTILAVSLSDVPAYMI-VTFKTVPLAISIGFLHGLVFL 819
QY 1259 FVILS-YVG 1266
Db 820 PLMLSVFVG 828

RESULT 53

US-10-415-934-9

; Sequence 9, Application US/10415934

; Publication No. US20040030099A1

; GENERAL INFORMATION:

; APPLICANT: Bayer AG

; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN

; FILE REFERENCE: LIO204 Foreign Countries

; CURRENT APPLICATION NUMBER: US/10/415,934

; CURRENT FILING DATE: 2003-05-05

; PRIOR APPLICATION NUMBER: US 60/245,565

; PRIOR FILING DATE: 2000-11-06

; PRIOR APPLICATION NUMBER: US 60/245,564

; PRIOR FILING DATE: 2000-11-06

; PRIOR APPLICATION NUMBER: US 60/245,572

; PRIOR FILING DATE: 2000-11-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 1061

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-10-415-934-9

Query Match

Best Local Similarity 7.7%; Score 531; DB 4; Length 1061;

Matches 185; Conservative 164; Mismatches 322; Indels 180; Gaps 27;

QY 492 QYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKDGTAALSCWAD----- 545
Db 5 EVFQELRLDNIIQIATTYDGDYTYKDN-----CARWNECPENDILNLDALMDIEAQ 61
QY 546 -----YGAPVFPFLAIGYKGYKGYSEAEALIMTFLSNYPAGDPRL----- 586
Db 62 LNLTFPMFNPVTDHLPF-VFPGTK-----LTDNYIVISVPAIQLVVFPV 107
QY 587 -----AQAKLWEAEFLBENRAFQRMAGMPQ---VTFTAERSLEBDEINKTTAEDLP 635
Db 108 TADTKRQDAKGAWEETFLRVVGNAN--SQQFKHISVSYPASRTLDHELEKNTKTVPY 165
QY 636 FATSIV--IFLYISALGSYSWSRVMVDSKATIGLGVAVVLGVAVMAAMGFFSYLGR 693
Db 166 FSTFLMLGLFSITCMGD-----AVRSKPFGLMGVNSAIMATLAFLAMGYCGT- 217

QY 694 SSLVILQVVPFLVLSVGADNIFIFVLVEYQRLPRRPGEPREVHIGRALGRVAPSMILCSLS 753
Db 218 EFIGINLAAPFLMIGIDDTFVMLAGWRTKAK--MPVAERMGLAMWSEAAVITITSVT 275
QY 754 EAICPFLGALTMPAVRTALTGLAVILDFFLOHSAFVALLSLDKRQBSRLDVC-CC 812
Db 276 DFISFLIGIISPRSVRICTYSVPAVCFTFLWHITFPFAACMAISGYRERKNLHSIFGCR 335
QY 813 VKPQELP-----PPQGEGLLGLGFFOKAYAPFLHLHWITRGV 848
Db 336 VQPMVAIKEKNFLYKATMAGIDANDPDNDINDKDMLMAFFKDKMAAVINKNCKAI 395
QY 849 VLLFLALFGVSLYSCHISVGLDQELALPKDSYLDYFLFLNRYFE-----VGAP 899
Db 396 IILAFASYLVGACYGITQIKGLERRKLSREDSYSVEFPDREDDYVREPPYRMQVLIAGP 455
QY 900 VYFVTTLGVNPSSEAGMNAICSSAGCNPFSTQKQYATEPPEQSYLAIPASSWDDFD 959
Db 456 L-----NYSDDLVPQEV-----ENLTSTLEHTSYVTSRRY-----TESWLSRPLS 495
QY 960 MLTPSSCCRLYISGPNKDKFCPSTVNSLCKNCKMSITMGSVRPSVEQFHKYLPWFLNDR 1019
Db 496 FLERN-----ELNVTVDDEQTFIDAVKEH--WLPQGN 527
QY 1020 PNIKCPKGLAAYSTVNLTSQGLASRFMAYHKPKNSQDYTEALRAARELANITA 1078
Db 528 P-----FSLDVRNEDETOIASRFLIQAVNITDTHKEKEMVR----- 565
QY 1079 DLKVPCTDPAPEVFPYITNVEYQYLILPEGLFMLSCLVPTTAVSCLLGLDLR 1136
Db 566 DLQKCKDSPLNASIFHPYP---VFPDQFELVRPVSLOQAMVIGAITMIISFVFIP-NIL 621
QY 1137 SGLNLLSIVMLVDVTGFMALWDISYNAVSLINLVSAGMSVRFVSHITSPALSTKPT 1196
Db 622 CSLWAPSVISIELGVAGVMAWLDVNLDSIMINIMCIGSFVDFAHICYTMSKKKS 691
QY 1197 WLERAKEATISMGSAAVAGVAMTNLPGILVLGLAKAQLIQIFPRLNLLITLLGLHGLV 1256
Db 682 PKARVREALHSLGLPIQSSSTIL-GIVALLAQSYIF-LVFFKMWFLVIFFGAMGLF 739
QY 1257 FLPVILSVGP 1267
Db 740 LLPVILLSLFGP 750

RESULT 54

US-11-097-143-23451

; Sequence 23451, Application US/11097143

; Publication No. US20050208558A1

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; FILE REFERENCE: CL000728

; CURRENT APPLICATION NUMBER: US/11/097,143

; PRIOR APPLICATION NUMBER: 2005-04-04

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

Db 400 TTSGSVSSAYSPFPSTSTINDILGKSEVSLKNIIILGYMFLIYVAVTL---IQW-RDP 455
Qy 662 VDSKATLGLGGVAVVLGAVMAAGFPFYLGRSLVILQVVPFLVLSVGADNIPFVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLFCALLGIPFNASTQIVPFLALGLGVQDMFLTHY 515
Qy 722 QRLPRRPGEBREHIGRALGRVAPSMCLLSLSAICFFLGAITPMPAVRTFALTSGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGVLSLASLNCVMFLAAALLPAPAFVFCQAAILLL 573
Qy 782 LDFLLQMSAFVALLSLDSKROEASRLDVCCVPOELPP----- 820
Db 574 FNLGSILLVFPAMISLDRRSARADLLCLMPESPPLPKKIPERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGEL-----LLGPFQKAYAPPLHMITRGVLLFLALFGVSLYSMC 865
Db 634 TROPLOPDVSENVTKCCLSVLTKWAKQYAFFIMRPAVKVITSMALIAVILTSVWGAT 693
Qy 866 HISVGLDQELALPKDSVLLDYFLNAR---YFVAGAPVYVVTTLGYNFSEAGNNAICSS 922
Db 694 KVRDGLDLTDIVPENT---DEHEFLSRQEKYFGF-----YN-----MYAVTQ- 732
Qy 923 AGCNFSF--TKIQVATEPPEOSYLAIP-----ASSWVDDPIDWL----- 961
Db 733 ---GNFEYPTNQKLLY---EYHQ-FVRIPNIIKNDNGGLTKFWLSLFRDMLLDLQVAFDK 786
Qy 962 -TPSSCCRLYISGPNKDKFCPTSVNSLNCIKNCMSITMGSVRPSVBO----- 1007
Db 787 EVASGCI-----TOEYWCNKASDE--GILAYKLMVQTGHVDNPDKSLITAGHRLVDKD 838
Qy 1008 -----FKHYL--PWFNLDRNPNTKCPKGGI-----AAYSTSVNLTSDBGVLA SRFM 1050
Db 839 GIINPRAFYNSAWATNDALAYGASQGNLKPQOPQRWIHSPEDVHLEIKKSSPLIYVQLP 898
Qy 1051 AVHKPLKNSQDYTEALRAARELANITADLRKVPQTDPAFEPVPIYITVFFHOYTILP 1110
Db 899 FYLSGLSDYXSIKTLRSVRDLCLKYEA--KGLPN-----FPGIPFLFWEQYLYLRT 949
Qy 1111 EGLFMLSCLVPTTFAVSCLLGLDLSGLNLSIVMLVDTVTGFMALWDISVNAVSLIN 1170
Db 950 SULLALACALAAVP--TAVVILLNNAVAVLTALATVQLLGVWALLGVKLSAMPVIL 1008
Qy 1171 LVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVIGLA 1230
Db 1009 LVLAIGRGVHTVHLCLGFVTSIGC---KRRASLALES-VLAPV---VHGALAAALA 1059
Qy 1231 KAQL-----TQIFFRNLILTLGLHGLVFLPVILSYGCP---DVNPALALEONRAE 1281
Db 1060 ASMLAASECGFVARLFLRLDLDIVFLGLIDGLLFFPILVLSILGPAAEVRPI----- 1110
Qy 1282 EAVAAMVWASCPNHRPSVST 1301
Db 1111 -----EHPERLST 1118

RESULT 56

US-09-754-032-4

; Sequence 4, Application US/09754032

; Publication No. US20030148388A1

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GOODRICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,032
; FILING DATE: 03-Jan-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
; US-09-754-032-4

Query Match 7.6%; Score 526.5; DB 3; Length 1311;
Best Local Similarity 22.2%; Pred. No. 9.3e-36;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;
Qy 336 LLGQPFQCGTGWASWPLTILVILVSVVLAAGLVFTLTDPVFLNSAPNSQARSEKA 395
Db 72 ILGCLFQDAGKLP--VAIILSTFCV-----GLKSAQIHRVDQLWQEGRLAEALK 124
Qy 396 FHDQHFQGP-FTNVOVILTAPNKRSSRYDLSLLGPKNPGSILDLLELLELLELLEL 454
Db 125 YTAQALGEADSTHQLVI---QTAKDPVSLIHP---GAL-----LEHL 162
Qy 455 QWMSPEAQONI-----SLQDICYAPLNPD-----NTSLYD-----CCINSILQYFQNN 497
Db 163 KVVHAATRVTHMYDIEMRLKOLCYSPSPIDPEGYHHIESIIDNVIPCAIITPLDCFWEG 222
Qy 498 RTLL-----LTTANOTLMGQTSQ--- 515
Db 223 SKLLGPDVPIYVPHLKHKLQWHLNPLVYVEVKKLKQFPPLSTIEAYMKRAGITSAYMK 282
Qy 516 ---VDWKDHFLYC-ANAPLTFKDGKTALALSCWADYGAPVPPF-----LAIGG----- 558
Db 283 KPCLDPTDP--HCPATAP--NKSGHI PDVAELSHGCTGFAAAAYMHPQLIVGGATRNS 339
Qy 559 ----YKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLMEAEFLBENRPAQRMAQMFQVT 614
Db 340 TSALRKARLQTVVQLMGEREMVEYWDHYKHQVQWQEKAAAVLDQWQFAAEVRKI 399
Qy 615 FT-----AERSLEDEINRTAEDLPFATSYVIFLYISIALGYSYNSWRVM 661
Db 400 TTSGSVSSAYSFPFSTSTLNDILGKSEVSLKNIIILGYMFLIYVAVTL---IQW-RDP 455
Qy 662 VDSKATLGLGGVAVVLGAVMAAGFPFYLGRSLVILQVVPFLVLSVGADNIPFVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLFCALLGIPFNASTQIVPFLALGLGVQDMFLTHY 515
Qy 722 QRLPRRPGEBREHIGRALGRVAPSMCLLSLSAICFFLGAITPMPAVRTFALTSGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGVLSLASLNCVMFLAAALLPAPAFVFCQAAILLL 573
Qy 782 LDFLLQMSAFVALLSLDSKROEASRLDVCCVPOELPP----- 820
Db 574 FNLGSILLVFPAMISLDRRSARADLLCLMPESPPLPKKIPERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGEL-----LLGPFQKAYAPPLHMITRGVLLFLALFGVSLYSMC 865


```
QY 962 -TPSSCRLYISGPNKDKPCPSTVNSLNCIKNCMSITMGSVRSVEQ----- 1007
Db 787 EVASGCI-----TQEYCKNASDE-GILAYKLMVQTGHVDMPIDKSLITAGHRLVDKD 838
QY 1008 -----FHKYL-PWFLNDRPNIKCPKGGI-----AAYSTSNNLSDGQVLASRFM 1050
Db 839 GIINPRAFNYLSAWATNDALAGASQGNLKPQORWIHSPEVHLEIKSSPLIYQLP 898
QY 1051 AYHKPLKNSODYTEALRAARELANITADLRKVPGTDPAPFPPYITITNVFYEQYLTIPL 1110
Db 899 FYLSGLSDTXSIKTLIRSVRDCLKYEA--KGLPN-----PPSGIPFLFWEQYLYLRT 949
QY 1111 EGLFMLSCLIVPFAVSCLLGLDLSGLNLLSIIVMLVDTVGFMAWDISVNAVSLIN 1170
Db 950 SLLALACALAAVF--TAVMVLLNAAVAVLVTALATLVLLQGLVNAALLGVKLSAMPAYL 1008
QY 1171 LVSAGVMSVEFVSHITRSFAISPKPTWLERAKATISMGSAVPAGVAMTNLPGILVLGLA 1230
Db 1009 LVLAIGRGVHTVHLCGLFVTSIGC-----KRRASIALES-VLAPV-----VHGALAAALA 1059
QY 1231 KAQL-----IQIFFRNLNLTLLGLLHGLVFLPVLISYVGP--DVNPALALEQKRAE 1281
Db 1060 ASMLAASECGFVARLEFLRLLDIVFLGLIDGLLFPFVILSILGPAAEVRPI----- 1110
QY 1282 EAVAAMVWASCPNHPRSVST 1301
Db 1111 -----EHPERLST 1118

RESULT 58
US-10-369-493-5464
; Sequence 5464, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5464
; LENGTH: 1405
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5464

Query Match 7.4%; Score 514; DB 4; Length 1405;
Best Local Similarity 20.2%; Pred. No. 1.3e-34;
Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLSDKLSFSTHTLLGQFQGWTVV--ASWELTILVLSVIPVVALAAGLVFTLTDD 378
Db 100 RKATGNRYALYSRSLIQKLLFALGNVTVRNWSI--ILAVSMIFAVC-CYGLQYVHIETD 157
QY 379 PYELMSA-----PNSQARSEKAFHQHFGPFF-RTN-----QVILTAPNRSS 419
Db 158 IVKLWAQGRDLDEELNPLNPKAERNVTGDS--GPPLPRENLGGGVQVLIQTPYEBG 215
QY 420 YRYSLLGPKNPSGILDLDLLELLELQERLRHLQVMSPEQRNI-----SLQDICY 472
Db 216 --QDALAAGP-----LLKHVEIMKHIASFNVSVHGVDWSLSDICF 253
QY 473 APLNPD-----NTSLYD-----CC-INSLLQYFQNNRTL-----LLL 503
Db 254 KPAPPSVAADSAASSLGVDIVKIVPCITWITPDCFWEGSGLGPHPSLPKSSLGPLGMILL 313
```

```
QY 504 TANQTMGQTSQVDMKD-----HFLYCANAPLTF----- 532
Db 314 SS-----LSDGDMIRNSDDPPIAVIDEIHRSFNLGSHYTFPERAGVSHGYMDRPCIDPLDP 369
QY 533 -----KGTALALSCMADYCAPVPPFLAIGGYKGDYSEAE 568
Db 370 ECPPMANVFDVCPHIDRVREIAKKGTELEBEKKKSGYSPPDFL--GRKKREAGDQ 426
QY 569 ALI----- 571
Db 427 KMIHPAQPADSIPTIEDAVPAQVPVUSTAPIPTITTLSPHEARAAEKKKQKARELKDYC 486
QY 572 -----MTYSLNNYP-----AGDP 584
Db 487 KSYKSAFEWLKKNKDKWPEVMSENMYQNVDDYAAEMTGCCSGFASNVLNWPMEDMILGNP 546
QY 585 RLAAQ--AKL-----WEEAFLEB-MRA 602
Db 547 RRAKGGKLGADALQSFLVASPADVFLRFKQKPGRNSMKTGLDMDAMNETAAEQVLQA 606
QY 603 FORMMAGMFQVTTAERSL-----EDEINRTTAEDLPIFATS-----YIV 642
Db 607 WQR-----NFT--KSLYNHKANVEDGNERRTLH--PLASTSIADMLEEFCQNYTI 654
QY 643 IF-----LYISLALGSYSWSRVMVDSKATLG--LGGVAVVLGAVMAAMGFSSYLGIRSSL 696
Db 655 ILAGYALMLAYAVITQARFDNCLPATESSMGLALAGLVVTPASVAGLATWFGIEFNA 714
QY 697 VILQVVPFLVLSVGADNIFIVLEVQRLPRRGPPEVHIGALGHVASMLLCSSEAI 756
Db 715 ATTQIVPFLTIGIGVDNMFMLLHNYRDVVVKGAGHAEMAI--LMRETGMSILCTSINNLI 772
QY 757 CFPLGALTMPAVRTFALTSLGLAVILDFLLOMSAFVALLSLDSKROEASRLDVCCCV--- 813
Db 773 SFLTGTLLPIPALRSFCAQSSILLTFNFIAITITPAILIISIDLRKKAORRDLVCLCYGD 832
QY 814 -----KPOBLPPPGQGEGL-----LLGFFQKAYA 837
Db 833 TREESYSMISKPKIKRIIGAPSEASIMQQFDGITQAQMASSDDDPAPWSLSHSFIRYVI 892
QY 838 PELLHWITRGVVLLFLALFGVLSISMCHISVGLDOELALPKDSYLLDYLFLFNRYFEVG 897
Db 893 PFISKASKVAIIVGCCALLGASFIGMRQSTLGLDLGVLPHTAPAQFLRARDKYFSF- 951
QY 898 APVYVFTVLGYNFS--SEAGMNAICSSAGCNPFSFTQKIQVATEPEGSYLAIPASSWV 954
Db 952 YPM-FAVINGPNI DYAHQORQIDNTRQSIGSSKYVKNK--NEEPSKY-----WL 999
QY 955 DDFIDMLTPSSCRLYISGPNKDKF--CPSTV-----NSLNC-----LKNCMSI 996
Db 1000 GLMRDMLI--SIQRGFDBEVAKGSFNLTSGTVIGSNVSEDAHLAHLAMCSHGLFCAG- 1056
QY 997 TMGSVR-----PSVEQPHKYL-PWFLNDR-----PNIKCPKGLLAAYSTS 1035
Db 1057 RVGKIRLVDAAGIINSDFNYLTAWFNVDHMMYTVSQASPPFTP-----PKWELSKNHT 1112
QY 1036 VNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVRPY 1095
Db 1113 NFIPAAEPLAVSQIPFYLTGLTDTTAVIDAKDRSVERP-----TQOGLPNFPQ 1163
QY 1096 TITNVFYEQYLTIPLFGLFMLSCLIVPFAVSCLLGLDLSGLNLLSIIVMLVDTVGF 1155
Db 1164 GIAFTFWQYLPFLTGNLMAQISIIITISVFCVLSVLL-FNPMAALMVVICILGIMTCELAF 1222
QY 1156 MALWDISVNAVSLINLVSAVMSVFEVSHITRSPAISTKPTWLERAKATISMGSAVPAG 1215
Db 1223 MGLVGIKLPNSAVTLITAVGIGVEFTVHVVSFL-----TALGTRSORTSSAVDRVFPY 1277
QY 1216 V---AWTNLPGTLVLGLAKAQLIQIFFERLNLILLGLHGLVFLPVLISYVGP--DVN 1270
Db 1278 VIHGSFTLLGILMLGSEFEPFVVKYFFIVMTALICIGILINGILLPVLSSPPRRREIS 1337
QY 1271 P 1271
```

```
Db      1338 P 1338
RESULT 59
US-09-898-533-3
; Sequence 3, Application US/09898533
; Patent No. US20020106656A1
; GENERAL INFORMATION:
; APPLICANT: Drabkin, Robert M.
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR.
; TITLE OF INVENTION: PATCHED
; FILE REFERENCE: 93445-00004
; CURRENT APPLICATION NUMBER: US/09/898,533
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: US/09/268,140
; PRIOR FILING DATE: 2000-03-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-898-533-3
Query Match      6.9%; Score 474; DB 3; Length 1286;
Best Local Similarity 20.6%; Pred. No. 3.4e-31;
Matches 246; Conservative 176; Mismatches 465; Indels 306; Gaps 43;
QY      321 KGTSLDKLSFSTHLLGQFFQGWGTWASWPLTILVLSVIPVVALAAGLVTELTDPV 380
Db      43 KGKGRSRTAIYLSRVFSQSHLETGSSVQKHAGKLVFVAILVLSFTCVGLKSAQIHSKVH 102
QY      381 ELWSAPNSQARKEKAFHDQHG-PPERTNQVILTA---PNRSYRYDLSLLGPKNFSGIL 436
Db      103 QLWIOGGRLAEALVTKTIGEDSATHQLLIQTHDNASVLHPQALL-----A 153
QY      437 DLDLLELELEQLERHLHQLVQWSPQASQRTSLQDICVAPLNPDNTSLY-----DCC 486
Db      154 HLEVLVKATAVKVHLVDTE-W-----GLRDMCMNPSTPGFEGYIYEIQLRHLIPCS 204
QY      487 INSLOYFQN-----NRTLLLLLTAN-----QTLMGQTSQVMDKDH 521
Db      205 IITPLDCFWEGSQLGPESAVVPGNLNQLLMTTLNPASVMQVMKQKMSKISDFETV 264
QY      522 FLY-----C-----ANAPLTFKD-----GTALALSQWADYGPVPPF 553
Db      265 EQMKRAAIGSGYMEKPCNLPLNPNCPDTAPKNKSTQPPDVGNALSGGCVG-YAAKHMW 323
QY      554 ---LAIGGYK---GKDYSAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
Db      324 PEELIVGGRKRNHSHLRKAQALQSVVQLWTEKEMDYQMDQDNKVHHLGWTQEKAAEVLN 383
QY      591 LWEEAP---LEEMRAFORWAGMFWQVTFABRSLEDEINRTTAEDLPFATSIVIVFLYI 647
Db      384 AMQRFNSREVEQLLRKQSRATNYDIYVFSAAALDDILAKFSHPALSIVIGVAVTVLY- 442
QY      648 SLALGSVSSMSRYMVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSSILVILQVFPFLVL 707
Db      443 --AFCTLLRW-RDPVRGQSSVGAVGLLMCFSTAAGLGLSALLGIYFNAASTQVVFPLAL 499
QY      708 SVGADNIFIVLBYQLRPRRPGPREVHIGRALGRVAPSMMLCSLSEACIFFIGALTMP 767
Db      500 GLGVGDHIFMLTAAYAESNR--EOTKL-----ILKKVGPSILFSACSTAGSFFAAAFIPV 553
QY      768 AVTFTALTGLAVILDFLLQMSAFVALLSDSKQESRLDV-CCCV-----KPELPP- 820
Db      554 ALKVFCLOAAIVMCSNLAAALVFPAMISLDLRRRTAGRADIFCCCFVWMKQPKVAPPV 613
QY      821 -----PGQGEGL-----LGFFQKAYAPP 839
Db      614 LPLNNNGRGARHPKSCNNRVRPLPAQNPLEQPADIPGSSHSLASFSLATFAQHYPFP 673
QY      840 LLHWITRGVLLFLALFGVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAP 899
Db      674 LMRSWVFLTWGFLAALISSLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTRLF--GFY 731
QY      900 VYFVTTLGYNFSSEAGMNAICSSAGCNFSTQKIQYATEPEPQSYLAIP----- 949
Db      732 SMYAVTQ3-----NFEYPTQQQLLRDY-HDSFVRVPHVKNKNGGL 771
QY      950 ASSWDDFIDWLTPSSCCRLYISGPNKDKPCFSTVNSLCLKNC-----MS 995
Db      772 PDFWLLLFSEWL-----GNLQKI PDEYRDGRRLTKECWFPNASSDAILAYKLI 819
QY      996 ITMGSVRPSVEQ-----FHKYL-PWFLNDRNENKPCPKGGLAAYSTS 1035
Db      820 VQTGHVNDPVDKELVLTNRLVNSGGINQRAFNYVLSAWTND-----VFAYG-- 867
QY      1036 VNLTSQGQVLASRFMAYHKPLKNSQD-----YTEALRAARELANITAD 1079
Db      868 ---ASQKLYPEPRQYFHP--NEYDLKI PKSLPLVYAQMPFYLHGLTDTSQI-KTLIGH 921
QY      1080 LRKVPGTDPAFEV--FPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLGLDLS 1137
Db      922 IRDLSVKYEGFGLPNYPSGIPFIWEQYMT-LRSSLAMILACVLALVLSLLLSVWA 980
QY      1138 GLMLLSITVMILVDTVGPMALWDISYNAVSLINLVSAGMSVEFVSHITRSPFISTKPTM 1197
Db      981 AVLTVLSVLASLAQIFGANTLLGKLSIPAIVALLISVGMMLCF--NVLISLGFMTSVGN 1038
QY      1198 LERAEATISGMSAVFAGVAMTNLPGIILVGLAKAQLIOIPFFRNLNLTLLGLLHGLVP 1257
Db      1039 RQRVQLSQMSGLGLVHGMLTSGVAVFMLSTSPPEFVIRHFCWLLLVLCVGCACNSLIV 1098
QY      1258 LPVILSYVGPDVNPALALEQKRAEAAVAVMVASCNHPSPRSTADNIVNHS 1310
Db      1099 FPIILSMVGPEAE-LVPLE-----HPDRISTSPPLPVRSS 1132
RESULT 60
US-11-097-143-4068
; Sequence 4068, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4068
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: DROSOPHILA
```


Db 43 KKGARSRTAIVLRVSFQSHLETLGSSVQKHAGKVLFAVILVLSVFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHQHFG-PFRTNQVILTA---ENRSSYRYDSLLGPKNFSGIL 436
Db 103 QLWIOGGGLEAEALAVTQKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLHLQVWPEAQRNISLODI CYAPLNPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVLHYDTE-W-----GLRDMCNMPTSPFEGYIYIEQILRHLIPCS 204
QY 487 INSLLOQYFQN-----NRTLILLTAN-----QTLMGQTSQVDWDKH 521
Db 205 IITPLDCFWEGSOLLGPESAVVPIGLNQLRLLTTLNPASVQMVKQKMESEKISDFETV 264
QY 522 FLYCANAPITFKDGTALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db 265 EQMKRAAIA---SGYMKPCPLNPLNCPDTPAKNKNSTQPPDVG-----ILSGGCGYA 317
QY 561 GKDYSEAEALIMTFLSNYPAGDPLAQA-----589
Db 318 AKHMHWPPEELIVGGAKRN-RSGHLRKAQALQSVMQVQMLTEKEMYDQMDNYKVHHLGWTOE 376
QY 590 -----KLWEEAF---LEEMRAFQREMAQMFQVTFPTAERSLEDEINRTTAEPLIPATSY 640
Db 377 KAEVUNAWQRFNSREVEQLLRKQSRITATNYDIYFSSAALDDIILAKFSPSALSIVIGV 436
QY 641 IVIFLYISALGSYSWSRVMVDSKATLGLGVAVVGLGAVMAAGMFFSYLGRSSILVILQ 700
Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAGLGLSALLGIVFNAASTQ 492
QY 701 VVPLVLSVGAADNIFIVLEVQRLPRPGEPREHVHIGRALGRVAVSMMLCSISEAICPFL 760
Db 493 VVPLFALGLGVHDHIFMLTAAYAESNR--EOTKL-----ILKVGPSILFSACSTAGSFFA 546
QY 761 GALTMPAVRTALTSLGLAVILDFLQMSAFALLSLDSKROEASRLDV-CCCV-----K 814
Db 547 AAPIPVPAUKVFLQAAIWMCSNLAALLVFPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606
QY 815 PQELPP-----PQOGBGL-----LLGFF 832
Db 607 PKVAPVPLVNNNGGARHPKSCNNRVALPAQNPLLEQRADIFGSSHSLSASPSLATFA 666
QY 833 QKAYAPFLHWTGRGVVLLFALFGVLSYMSCHISVGLDQBLAPKDSYLLDYFLFLNR 892
Db 667 PQHYTPFLMRSWVKFLTMVGMFLAALISLYASTRLQDGLDIIILVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVVTGLVNFSSBAGMAICSSAGCNPFSTQKIQYATEFPEQSYLAIP---- 949
Db 727 LP--GFYSMYAVTQG-----NFEYPTQOQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWDDFDIDLTPSSCCRLYISGPNKDKFCPSVNSLNLKNC----- 993
Db 765 KNDNGGLPDFMILLSEML-----GNLQKIFDEERYDGLTEKCPFNASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db 813 ILAYKLIVQTHGVNDPVDKELVLTNRVNSDGIINQAFYNYLSAW-----ATNASSPTL 868
QY 1029 LAA--YSTSVNLTSQGVLASRPMAYKPLNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQGLYPEPQVPHOP--NEYDLKIPKSLPLVYAQMPFYLHGLTDTTS 926
QY 1071 ELAANITADLRKVPQTDPAFV--PEYTTITNVEQYLTILPEGLMSLCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDLVKYEGFGPNYPSGIPFIFWEQYMT-LRSSGLAMILACVLLAALVLV 984
QY 1129 LLGLDLRSGLMLLSILMVLVDVTFGMALWDISVNAVSLINLSVAGVMSVBFVSHITS 1188
Db 985 SILLLSVWRAVLVLSLASLAQIFCAWTLGLGLKLSAIPAVILILSVGMMLCF--NVLS 1042
QY 1189 FAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQIQLIFFFLMLLITL 1248
Db 1043 LGFMTSVGNQRVRVQSLMQMSLGPLVHGMLTSGVAVFMSLSTSPFEPVIRHFCWLLILVLC 1102

QY 1249 LGLLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVMVASCPNHPRSVSTADNIYVN 1308
Db 1103 VGACNSLLVFPILLSMVGPARE-LVPLE-----HPDRISTPSPLVR 1143
QY 1309 HS 1310
Db 1144 SS 1145

RESULT 63
US-10-835-517-48
; Sequence 48, Application US/10835517
; Publication No. US20050004028A1
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; McMahon, Andrew P.
; Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/835,517
; FILING DATE: 28-Apr-2004
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,771
; FILING DATE: 20-OCT-1997
; APPLICATION NUMBER: US 08/462,386
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 04-MAY-1995
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMV-006.11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-835-517-48

Query Match 6.7%; Score 464.5; DB 5; Length 1299;
Best Local Similarity 20.4%; Pred. No. 2.3e-30;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSPTHTLLGQFFQCGMTWVASWPLTILVLSVIVWALAAAGLVFTELTDPV 380
Db 43 KKGARSRTAIVLRVSFQSHLETLGSSVQKHAGKVLFAVILVLSVFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHQHFG-PFRTNQVILTA---ENRSSYRYDSLLGPKNFSGIL 436

Db 103 QLWIOGGGLEAEALVAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
Qy 437 DLDLLELLELQERLRLHQLVWSPQAQRNLSLDICYPALNPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVLHYDTE-W-----GLRDMCMPSTPSFGIYIEQILRLHLPSCS 204
Qy 487 INSLQYFQN-----NRTLALLLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFEGSQQLGPESAVVPIGLNQRLWTLTLPASVMQMKQKSEKISIDFETV 264
Qy 522 FLYCANAPLTFKDGKTALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db 265 EQVMKEAATA---SGYMERPCPLNPLNCPDTPAPKNKSTQPPDVG-----ILSGGCGYA 317
Qy 561 GKDYSEAEALIMTFSLNNTYPAGDPRLAQA-----589
Db 318 AKHMHVPEELIVGGAKRN-RSGHLRKAQALQSVMQTEKMYDQMDQNKVYKHHLGWTOE 376
Qy 590 -----KLWEEAF---LEEMRAFOREMAGMFQVTFTAERSLEDEINRTTAEDLPFATSY 640
Db 377 KAAEVUNAWQRFNSREVEQLLRKQSRITATNYDIYVSSAALODILAKFSPHSAISVIGV 436
Qy 641 IIVFLYISALGSYSSWMSVMDSKATLGLGVAVVVLGAVMAAMGFYSVLGRSSILVIQ 700
Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGAGVLLMCFSTAGLGLSALLGIVFENASTQ 492
Qy 701 VPFPLVLSGADNIFTVLEYORLPRRPOGEPREHIGRALGRVAPSMMLCSLSEACFPPL 760
Db 493 VVPFLALGLGVDFHIFMLTAAEAESNR---EOTKL-----ILKKVGPSILFSACSTAGSPPA 546
Qy 761 GALTMPAVRTFALTSLGAVILDFLLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
Db 547 AAFIPVPAKVFCLQAAIWMCSMLAAALIVFPAMISLDLRRRTAGRADIFCCCFPPWKQ 606
Qy 815 PQELPP-----PQGEGL-----LLGFF 832
Db 607 PKVAPPVPLNNNGRGARHPKSCNNRVALPAQNPLLEQRADIPGSSHSLASFSLATFA 666
Qy 833 QKAYAPFLHWTGRVVLVLLFLAFGVLSYMSCHISVGLDQELALPKDSYLDLYFLFLNR 892
Db 667 FQHYTFPLRWSWKFLTVNGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
Qy 893 YPEVGAIVVFTVLTYGVNFSSEAGMNAICSSAGCNFSTQKIQYATEPPEQSVLAP--- 949
Db 727 LP--GFYSYAVYTOG-----NFEYFTQOQLRDY-HDSFVRVPHVI 764
Qy 950 -----ASSWVDDFDLWLTSPSCCRLYISGNPKKFCPSTVNSLNCNLC----- 993
Db 765 KNDNGGLPDPFLLLFSEWL-----GNLQKIFDEYRDGRLTKECFPNASSDA 812
Qy 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRNPKCPKG 1028
Db 813 ILAYKLIIVQTHGVNDPVDKELVLTNRNVNSDGIINORAFNYLSAW-----ATNASSPTL 868
Qy 1029 LAA--YSTSVNLTSQVLSRPMAYHKPLKNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQGLKPEPEQYFHPQ--NEYDLKIPKSLPLVVAQMPFYLHGLTDS 926
Qy 1071 ELAANITADLRKVPDTPAFEV--FPYITINPYEQYLATLPEGLFMLSCLVPTPAVSC 1128
Db 927 QI-KTLIGHIRDLVSVKYEGFGLPNYPSPGIPFIWEQYMT-LRSSLAMILACVLLAALV 984
Qy 1129 LILGLDLRGLLNLISVIMLVDTVGFMAWDNISYNVSLINLVSAVGSVEFVSHITS 1188
Db 985 SLLLSVMAAVLVLSVLAASLAQIFGAMTLGLKLSAIPAVILISVGMMLCF--NVLIS 1042
Qy 1189 FAISTKPTWLERAKEATISGSAVFAGVAMTNPGLVLGLAKAQIQLIFFFLNLILTL 1248
Db 1043 LGFMTSVGNRQRRVQLSMQMSLGPLVHGMLTSGVAVFMLTSPFPFVIRHFCWLLLVLC 1102
Qy 1249 LGLHGLVFLPVLTLSYVGPDPNPALALEOKRAEEAAVAVMVASCPNHPRSVSTADNIYN 1308
Db 1103 VGACNSLLVFPILLSMVGPEAE-LVPLE-----HPRISTPSPLPVR 1143

Qy 1309 HS 1310
Db 1144 SS 1145
RESULT 64
US-10-890-776A-4809
; Sequence 4809, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 4809
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: Sequence of the Patched motif of gi 3892144
US-10-890-776A-4809
Query Match 6.4%; Score 442.5; DB 5; Length 906;
Best Local Similarity 20.0%; Pred. No. 1.1e-28;
Matches 202; Conservative 196; Mismatches 367; Indels 243; Gaps 39;
Qy 378 DPVELWSAPNSQARSEKAFHQHGFQFFRTNQVILTAPNRSRYRDSLLGPKNFSGILD 437
Db 5 DPSYVFTSDARWRREISVFENW-PL--DENKFLPGKSFEAKRFVNLIRAKDGGSIMR 61
Qy 438 LDLLELLEL-QERLRLHQLVWSPQAQRNLSLDICYPALNPNTSLYD--CCINSLQYF 494
Db 62 DNVLHEIELANQWIMNNSIPTDDLKFNLTQYDCLLS-----YDWCVCANEHIQ-- 110
Qy 495 QNRTLLILLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGKTALALSCWADYGAPVFPPL 554
Db 111 -----MLLRND--VNQILDL---HF-----P---RGGTK-----DTPVY--- 137
Qy 555 AIGYKGKDYSEAEALIMTFSLNNTYPAGDPRLAQA-----AKLWEE---AFLEEMRAP 603
Db 138 -LGGIFGD-----VQFF--QNGTSLDAKLTQLFYFLKQDQKQVVEEYSSKFSYALETF 186
Qy 604 QRQM--AGMFQVTFTAERSLEDEINRTTAEDLPFATSYVIFLYI---SLALGSYSSWS 658
Db 187 LNQVYSSDVITLSFAHYQSLEDGLDENAKAFVFNFWVSFFVLAMVALVSSFTLKSSSATK 246
Qy 659 RVWVDSKATLGLGVAVVVLGAVMAAMGFYSVLGRSSILVILQVVPPLV-----707
Db 247 IDWISKKPMAAAGMFPSTVLSIISAFGLFIILGVRYN-VINTIIIFLIIGFQSFQKMEK 305
Qy 708 ----SVGADNIFVLEYORLPRRPOGEPREHIGRALGRVAPSMMLCSLSEACFPFLGAL 763

```
Db 306 NQPAIGIDMFLLNACWDTSKSLVPE--RMSKTLSHAGVAVTITNVTDVMSFAIGCI 363
Qy 764 TPAVTRTALTSLGLAVILDLLQMGAFVALLSL-----797
Db 364 TDLPGIQFFCIYACVSFAFSYFQLAFFSGAMAIMGEVEREKHCLPFYRTTQLVDISKM 423
Qy 798 -----DSKRQE-----ASRLDVCC--CVKQELPPPG--822
Db 424 NEBADSQLOQIKRSASPAPFNFLSSNSGFSFSDSFSKTKTIPAEFAWKEQQSPNSSL 483
Qy 823 -----QEGSLLGFFQKAYAPFLHMTIRGVVLLLELALFGVSLYSMCHISVGLDQEL 875
Db 484 SKSKDREKDRIVHFIKIGYGFILNSVRIPSGLIFFVYLAIAMYGCVNFRBGLNPGN 543
Qy 876 ALPKDSYLLDYFLNRYFEVGPVYFVTTGLGYNFSEAGMNAICSSAGCANNFSTFKTIQ 935
Db 544 LVTDHYIAKYFSDIKHFWRIGQAHLVAVLNPNTLISENNEL-----LK 589
Qy 936 YATEPPEQSYLAIPASSWDDDFIDWLTTPSSCCRLYISGPNKDKFCPSTVNSLCLNCKMS 995
Db 590 VVSFAFNTQY--TLGREGTVFFLLLEYLN-----YLSLNAE-----VE 625
Qy 996 ITMGSVRPSVEQPHKYLPHFLNDRPNIKCPKGLAAYSTSVMLT--SDGOVLASRPMAYHK 1054
Db 626 DTERLWTKLNSWLKY-----TGSTQWASNLKINKTDGSGFQAFRQIALK 671
Qy 1055 PLKNSQDYTEALRAARELAN-----ITADLRKVPGTDPAPFVFPYITNV-----1100
Db 672 NFVEPNHDKHAQALLADIADHPFNVVVHVSEVSGNRKIL--NDFISSHSCYAKNIPKLA 730
Qy 1101 --FYEQYLITLP---EGLFMLSCLVPTPAVSCLLGLDLRSLGLMLLSIVMLVDVTG 1155
Db 731 PFPADQYLILPATIQNVVISLLCM-----AVVSFLLVPSLPSGFFVFSIVSINIGVGY 786
Qy 1156 MALWDSYNVAVSLINLVASGVMSVEFVSHITSPALSTPTWLERAKEATISMGSAVPA 1215
Db 787 MTLGVNLDAVMSISITMSIGFAVDLSAHIYAFVTSBGDT--KORVIGALETGLWPIFOG 845
Qy 1216 VAMTNLPGLVLGLAKAQIQFFRLNLILTLGLLHGLVFLPILS 1263
Db 846 ASST--IAGSILYTVDAYII--LVFFKTIWLTMLIGNHGLFPIPIPLS 891
```

RESULT 65

```
US-08-954-701A-6
; Sequence 6, Application US/08954701A
; Publication No. US20030032085A1
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
```

```
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-701A-6
```

```
Query Match 6.3%; Score 438.5; DB 2; Length 1285;
Best Local Similarity 20.0%; Pred. No. 4.3e-28;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

Qy 321 KCTSLSDKLSFSTHTLLGQFPQGWGTWVASWPLTLLVLVPIPVVALAAGLVFTELTDPV 380
Db 43 KKGKARGSTAIYLRSVFQSHLETGLSSVQKHAGKVLFAVLVLSTFCVGLSKAQIHSKVH 102
Qy 381 ELWSAPNSQARSEKAFHQHFG--PFPRTNOVILTA---ENRSSYRVDSLLGLPKNFSGIL 436
Db 103 QLWIOEGGRLEARLAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
Qy 437 DLDLLELELEQLERLRLHQVNSPEAQRNISLQDICYAPLNPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHLVDYE--W-----GLRDMCNMFPSTPSEGIYIIEQILRHLPIS 204
Qy 487 INSLIQYQFN-----NRTLLLLLTAN-----QTLMGQTSQVDMKDH 521
Db 205 IITPLDCWEGSQLLGPESAVVIPGLNQRLLTTLNPASVMQVMQKMSSEKISDFETV 264
Qy 522 FLY-----C-----ANAPLTFKD-----GTALALSCHADYAGVPPFP 553
Db 265 EGYMKRAAIGSGYMKPCPLNPLNCPDTPAPKNKSTQPPDVGAILSGGCY--YAAKMHM 323
Qy 554 ---LAIGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----590
Db 324 PEELIVGGRKRNRSGLRKAQALQSVQLMTEKEMVDQMDNYKVHHLGWTQEKAEVLN 383
Qy 591 LWEEAF---LEEMRAFQRMAGMFQVTTAERSLSDEINRTTAEDLPFATSYVIFLYI 647
Db 384 AMQNFREVEQLLRKQSRITATNYDIYVFSAAALDILAKFSGHPSALSIVIGVAVTVLY- 442
Qy 648 SLALGSYSWSRVWDSKATLGLGVAVVILGAVMAAGFFSYLGI-----692
Db 443 --AFCTLLRW--RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
Qy 693 --RSSLVI---LQVVPFLVLSVGDNIPIFVLEYQLPRRPGEPREVHIGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVPFLALGLGVDFIPI-----VGPSI 533
Qy 748 LLCSLSEACFPFLGALTMPMPAVRTPALTSGLAVILDFLLQMSAFVALLSLSKQREASRL 807
Db 534 LFSACSTAGSPFAAAPIVPALKVFCLQAAIVMCSNLAALLVFPAMISLDLRRRTAGRA 593
Qy 808 DV-CCCV-----KQELPP-----PQ 823
Db 594 DIFCCCFPVWKEQPKVAPPVPLPLNNNGRGARHPKSCNNNRVPLPAQNPLLSQRADIPGS 653
Qy 824 BEGL-----LLGFFQKAYAPFLHMTIRGVVLLLELALFGVSLYSMCHISVGLDQELALPK 879
Db 654 SHSLASPSLATPAFOHYTTPFLMRSVWKFLTWGFLAALISSLYASTRLQDGLDIDLVPK 713
Qy 880 DSYLLDYFLNRYFEVGPVYFVTTGLGYNFSSSEAGMNAICSSAGCANNFSTFKTIQYATE 939
Db 714 DSNHKKFLDAQTRLF--GFYSMAVTVQ-----NFEYPTQQQLLRD 752
Qy 940 PPEQSYLAIP-----ASSWDDDFIDWLTTPSSCCRLYISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNKNDGGLPFWLLLFSEML-----GNLQKIFDEYRDGRL 798
```

Qy	990	LKNC-----MSITGVSVRPSVEQ-----PHXYL-PWF 1015
Db	799	TKECWPNASSDAILAYKLI VOTGHVDNPVDKELVLTNRILVNSDGIINQRAFYNYLSAWA 858
Qy	1016	LNDRPNIKCPKGGLAAVSTSVNUTSDGOVLASRFMAYHKPLNSQD----- 1061
Db	859	TND-----VFAYG-----ASQKLYPEPRQVPHQ--NEYDLKPKSLPLVYAQM 901
Qy	1062	--YTEALRAARELAANITADLRKVPGTDPAPFV--PFYTTNVFYEQYLITLPEGLFMLS 1117
Db	902	PFYUHGUTDSQI-KTULIGHIRDLSVKYEGFGLPNTPSGIPPFIEWEQYMT-LASSLAMIL 959
Qy	1118	LCLVPTFAVSCLLGLDLRGLNLLSIVMLVDVTGCFMALWDISYNAVSLINLVAQVM 1177
Db	960	ACVLLAALVLVSILLLSWMAAVILSVLASLAQIFGAMTLLGKLSAIPAVAILIUSVMG 1019
Qy	1178	SVEFVSHITSPALSTKPTWLERAKEATISNGSAVFAGVAMTNLPGLTVLGLAKAQLIQI 1237
Db	1020	MLCF--NVLISLGFWMTSGVNRQRVQUSMQMSLGLPLVHGMLTSGVAVFMLSTSPFPEVIR 1077
Qy	1238	FFPRLNLIITLGLLHGLVFLFVLISXVGPDPNPALALEQKRAEBAVAAVWVASCNPHPS 1297
Db	1078	HFCWLLLVLCVGACNSLLVPFILLSMVGPEAE-LVPLE-----HPD 1118
Qy	1298	RVSTADNIYVNHVS 1310
Db	1119	RISTSPPLPVRSS 1131

Db 1020 MLCP--NVLSLGFMTSGVNRQRVQLSMQNSLGLVHGMLTSGVAVFMLSTSPFFEVIR 1077
QY 1238 FFPLNLITLLGLHGLVFLVSLVGVDPDVPNPALEOKRAEEAFAVAVMVASCPNHP 1297
Db 1078 HFCWLLVVLVCGACNSLLVFPILLSVMGPEAE-LVPLE-----HPD 1118
QY 1298 RVSTADNIYVNH 1310
Db 1119 RISTPSPVPRSS 1131
RESULT 68
US-10-493-493-5410
; Sequence 5410, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052) B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5410
; LENGTH: 714
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5410
Query Match 5.5%; Score 379; DB 4; Length 714;
Best Local Similarity 23.0%; Pred. No. 2.6e-23; Mismatches 299; Indels 160; Gaps 25;
Matches 173; Conservative 119;
QY 545 DYGAFFPFLATGGYKGDYBAEALIMTFSLN--NYP----AGDPRLAQ--AKLWEERAF 596
Db 58 DYAEH-----TGCGSG-----FASVNLWNPEDMILGNPRRAKGGKLSGADA 100
QY 597 LEEMRAFORMAGMFOVTTAERSLEDEINRTAEDLPATSYIVIFLY----- 646
Db 101 LQ----FVFLVSPADVFLFRKQSLSE-----TLETLSLYEEHCHLHRRHPLLCRSNSV 151
QY 647 ----ISLALGSYSWSRWV-----DSKATLGLGVAVVLGAVMAAGPFSL 690
Db 152 NSTIPSSLLDMLSCWMLLLILLDTTAFRLQHONGFALAGVLVVTTFASVAGIGLATWF 211
QY 691 GIRSSLVILQVFPFLVSLVGADNIFVLEYQRLPRRPGPREVHIGRALGRVAPSMLLC 750
Db 212 GIFENATTOIVFLTLGIGVDNFMFLHNYRDVVKLAGHAEMAI--LMRETMGSILCT 269
QY 751 SLSEACIFFLALTPMPAVRTFALTGLAVILDFLOMSAFVALLSDSKRQPSRLDVC 810
Db 270 SINILSFLGTLLPILPALRSFCAQSSILLTFNFIAITLYPAIISIDLERKKAQRDFD 329
QY 811 CCVK-POELPPPGGEG-----LLGFFOKAYAPFLLHMTTRGVLLLLFL 854
Db 330 TREESTSEASIMQOFGDITOAQWASSDDPAPWSLHSFIRYIYPTFKPASKVAIIIGCC 389
QY 855 ALFGVSLYMSCHISVGLDQELALPKDSYLLDYFLNRYPEVGAPVYVVTTLGVNFS--- 911
Db 390 ALLGASFVGMRSQTLGLGLGVLPHTAPAQFLRARDKYFSF-YPM-FAVIKGNIDYAH 447
QY 912 SEAGNAICSSAGCNFSFTQKIQYATEFPQSYLAIPASSWDDFTDMLTPSSCCRLXI 971
Db 448 QQRQIDNRYQSIGSSKVIKNK-----NEEPSEK-----WLGIMRDLI--SIQGF 494
QY 972 SGPNNKDF-----CPSTVNSLNCNCSITMGSVRPSVEQPHK--YLPFWFLNDRPNI 1022

Db 495 EEVAKGSFNLTSGTVIGSNVSEDARLAHALMCHSGSLFCACRIGKIRLIPFYL----- 549
QY 1023 KCPKGLAAYSISVNLTSQGVLASRFMAYHKLKNSQDYTEALRAARELANITADLRK 1082
Db 550 -----GLTDTAVTDAIKDIRSVCFR----- 571
QY 1083 VPCTDAFEVFPVTTITNVFVEQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLNL 1142
Db 572 ----TDQGLPNFPQGIATFWEQYLFGLTNLMOAISITISVFCVISVLL--FNPWALMVV 627
QY 1143 LSIWMLVDTVGFMAWLDISYNAVSLINLVSVMGSEVFEVSHITRSFAISTKPTWLERAK 1202
Db 628 CILGIMTCELAFMGLVGLIKLNPVSAVTLITAVGIGVEFTVHVVSFL-----TALGTRS 682
QY 1203 EATISMGSAVAGV---AMTNLPGILVGLA 1230
Db 683 QRTSSAVDRVFPVPIHSGSFSTLIGILMLGFS 713
RESULT 69
US-10-060-756A-6
; Sequence 6, Application US/10060756A
; Publication No. US20030046717A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/060,756A
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4804
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 6
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-756A-6
Query Match 5.1%; Score 351.5; DB 4; Length 767;
Best Local Similarity 20.5%; Pred. No. 7.3e-21; Mismatches 285; Indels 189; Gaps 31;
Matches 158; Conservative 137;
QY 334 HT-----LLGQFFQCGMTWVASWPLTILVSLVPIVVALAAGLVF--TELTDFVELWSAP 386
Db 115 HTDCLLEGLSRTFQMLGWQVGAHPFLLAPLMLTAALGTGFLYLPKDEEDLEEHVTPV 174
QY 387 NSQARSEKAFHDPQHPFPPTNQVILTAPNRS-----YRDSLLGLPKNFSGL 436
Db 175 GSPAKAERFVQGH-----FTTNDYSRFSASRRSTEANFVSLVSVSYSDSLDDPATFAEVS 230
QY 437 DLDDLELLELQELERLHLQVWSPQAQRNLSLODIC--YAPLNPNTSLYDCCI--NSLLQ 492
Db 231 KLDGAVQDLRV-----AREKGSQYQVQVCARYAL-----CVPPNPILY 270
QY 493 YFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALALSMAVYGA-PVF 551
Db 271 AWQVNTKLN-----SSISFPAY-----NHRHPLY 296

```
QY 552 PFLAIGGY-----KGKDYSEAEALIMTFSLNNYPAGDPRL-AQAKLWEEAFLEEMRAF 603
Db 297 LTGFFGGYILGSLGMLQGLLRKAMRLLYLK---TEDPEYDVQSKQMLTHLLDQFTNI 353
QY 604 QRMA--GMFQVFTTA-ERSLEDEINRTTAEDLPIEATSVIVILFYLISALGSSWSRV 660
Db 354 KNILALKKIEVVHFTSLRQLEFATSVTV--IPVFLAYILILFVNTSCFRFD-----406
QY 661 MVDKATLGLGGVAVVLGAVMAAGPFSYLGRSSILVQVVPFLVLSGADNIFIFVLE 720
Db 407 CIRNKMCAAFGVISAFVAVSGFGLLHIGV-PFVLIIVANSFPLILGVGDDMFIMISA 465
QY 721 YQRLPRPEPREVHIGRALGRVAPSMCLCSSEALCFPLGALTMPAVRTALTSLGLAV 780
Db 466 WHK--TNLADDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSVQCFTYGMTL 523
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELP-----PPGOG 824
Db 524 LFCYFYNITCFGAFMALDKREV-----VLCWLKKAADPKWPSFKKCFPPGSPVDEHG 578
QY 825 EGL--LLGFFQKAYAPFLHWTGRVVLFLFALFGVSLYSMCHISVGLDQBLALPKDSY 882
Db 579 TDIHPISLFRDYFGFPLTRSESKYFVVIYLYIISIIYGCFFHQBEGDLNRLASDDSY 638
QY 883 LLDYFLFLNRYF-EVGAPVYFVTLTYGNFSSBAGMNAICSSAGCNFSTQKIYATEPP 941
Db 639 ITPYFNVEENYFSDYGRVMVITKVDYWDK-----DVRQKLENTKIF 683
QY 942 EQSYLAIP--ASSWVDDFIDMLTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITWG 999
Db 684 EKNVYVDKNTLTFWLDAYVQ-----YLGNSQD---PNEKNT-----717
QY 1000 SVRPSVEQFHKYLPMFLNDRPNIKCPKGLAAYSTSVNLTSDGVLASR 1048
Db 718 -----FMNIPDFLSNFPN-----FQHDINISSNEIISR 748

RESULT 70
US-10-890-776A-6
; Sequence 6, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 6
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-890-776A-6
```

```
Query Match 5.1%; Score 351.5; DB 5; Length 767;
Best Local Similarity 20.5%; Pred. No. 7.3e-21;
Matches 158; Conservative 137; Mismatches 285; Indels 189; Gaps 31;

QY 334 HT-----LLGQFQCGTWWASWPLTILVLSVIVPVALAAGLVF--TELTTPDVELWSAP 386
Db 115 HTDCEGLLSRTFQWLGWQVGAHPWIFLLAPLMLTAALGTGLYLPKBEEDLEEHYTPV 174
QY 387 NSOARSEKAFHDQHFGRPTNOVILTAPNRSS-----YRDSLLLGKPNKSGIL 436
Db 175 GSPAKAERRFVQGH-----FTTNDSTRFSASRSTEAANFVSLVSVSDSLDLPATFAES 230
QY 437 LLDLLELELELERLHLQVWSPERQNRISLQDIC--YAPLNPDNSTSLYDCCI--NSLLQ 492
Db 231 KLDGAVQDLRV-----AREKGSQIQYQVCARYEAL-----CVPPNPILY 270
QY 493 YFQNRRTLLLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGTTALALSMDADYGA-PVF 551
Db 271 AMQVNTKTLNL-----SSISFPAY-----NHGRHPLY 296
QY 552 PFLAIGGY-----KGKDYSEAEALIMTFSLNNYPAGDPRL-AQAKLWEEAFLEEMRAF 603
Db 297 LTGFFGGYILGSLGMLQGLLRKAMRLLYLK---TEDPEYDVQSKQMLTHLLDQFTNI 353
QY 604 QRMA--GMFQVFTTA-ERSLEDEINRTTAEDLPIEATSVIVILFYLISALGSSWSRV 660
Db 354 KNILALKKIEVVHFTSLRQLEFATSVTV--IPVFLAYILILFVNTSCFRFD-----406
QY 661 MVDKATLGLGGVAVVLGAVMAAGPFSYLGRSSILVQVVPFLVLSGADNIFIFVLE 720
Db 407 CIRNKMCAAFGVISAFVAVSGFGLLHIGV-PFVLIIVANSFPLILGVGDDMFIMISA 465
QY 721 YQRLPRPEPREVHIGRALGRVAPSMCLCSSEALCFPLGALTMPAVRTALTSLGLAV 780
Db 466 WHK--TNLADDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSVQCFTYGMTL 523
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELP-----PPGOG 824
Db 524 LFCYFYNITCFGAFMALDKREV-----VLCWLKKAADPKWPSFKKCFPPGSPVDEHG 578
QY 825 EGL--LLGFFQKAYAPFLHWTGRVVLFLFALFGVSLYSMCHISVGLDQBLALPKDSY 882
Db 579 TDIHPISLFRDYFGFPLTRSESKYFVVIYLYIISIIYGCFFHQBEGDLNRLASDDSY 638
QY 883 LLDYFLFLNRYF-EVGAPVYFVTLTYGNFSSBAGMNAICSSAGCNFSTQKIYATEPP 941
Db 639 ITPYFNVEENYFSDYGRVMVITKVDYWDK-----DVRQKLENTKIF 683
QY 942 EQSYLAIP--ASSWVDDFIDMLTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITWG 999
Db 684 EKNVYVDKNTLTFWLDAYVQ-----YLGNSQD---PNEKNT-----717
QY 1000 SVRPSVEQFHKYLPMFLNDRPNIKCPKGLAAYSTSVNLTSDGVLASR 1048
Db 718 -----FMNIPDFLSNFPN-----FQHDINISSNEIISR 748

RESULT 71
US-10-094-749-1772
; Sequence 1772, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: TETSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
```

Db 550 ETNREELYDLLETLREL--SVTSKVKFIV-FNPSP-----VTMDRYASSL--GAPL 595

Qy 1116 LSLCLVPTTAV--SCLLGLDLRSGLNL--LSIYVILVDTVGFMAWDISNAVSLIN 1170

Db 596 HNSCISALFLLPFSAFVA---DSLINVMITLTVSVSEFVGIVGFWTLWKVELDCISVLC 651

Qy 1171 LVSAGVMSVEFVSHITRSPAIS---TKPTWLERAKEATISMSGSAVFAGVAMTN-----LP 1222

Db 652 LIYGINVTINDCAPMSAFVLGKDFTRTKWKNALV-----HGVAILQSYLCYIV 702

Qy 1223 GIDVLGLAKAQLIQIFFRNLNLTLLGLLHGLVFLPVILSYVGPVDPNPALEAEKRAEE 1282

Db 703 GLIPLAAVPSNL-TCTFLRCLFLIAFVTFHCFAILPVILTFLPPS-----KCKRKEK 754

RESULT 72

US-10-060-756A-10

; Sequence 10, Application US/10060756A

; Publication No. US20030046717A1

; GENERAL INFORMATION:

; APPLICANT: Zhang, Jian

; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN

; FILE REFERENCE: PB0177

; CURRENT APPLICATION NUMBER: US/10/060,756A

; CURRENT FILING DATE: 2002-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 09/864,761

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/327,898

; PRIOR FILING DATE: 2001-10-09

; NUMBER OF SEQ ID NOS: 4804

; SOFTWARE: Acomica Sequence Listing Engine

; SEQ ID NO 10

; LENGTH: 648

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-060-756A-10

Query Match 4.5%; Score 310.5; DB 4; Length 648;

Best Local Similarity 22.0%; Pred. No. 2.1e-17;

Matches 133; Conservative 107; Mismatches 238; Indels 127; Gaps 24;

Qy 334 HT-----LLGQFFQCGGTWVASWPLTILVSVIPVVALAAGLVF--TELTTDPVELWSAP 386

Db 115 HTDCLGSLLSRTFQWLGWQVGAHPWIFLLAPLMTAALGTGLYLPKDEEDLEEHYTPV 174

Qy 387 NSQARSEKAFHOGFGFPFRTNQVILTAPNRSS-----YRYDSLILGPKNFSGIL 436

Db 175 GSPAKAERRFVOGH-----FTTNDSYRFSASRRSTEANFVSLLVVSDSLDPAFAVS 230

Qy 437 DLDLLELELEQLRHLQVSPQAORNTSLQDIC--YAPLNPDNTSLYDCCI--NSLLQ 492

Db 231 KLQGAQDLRV-----AREKSGIQYQQVCARYAL-----CVPPNPILY 270

Qy 493 YFQNNRTLLLTANQTLMGOTSOVDKWDHFLYCANAPLTFKDTALALSCWADYGA--PVF 551

Db 271 AQVQNKTLNL-----SSISFPAY-----NNGRPL 296

Qy 552 PFLAIGY-----KQKDYSEAEALMTFSLNNYPAGDPRL-AQAKLWBEAFLEEMRAF 603

Db 297 LTGPFPGYIILGSLGMQGLLLRAKAKMLLYLK---TEDPEYDVQSKWLTHLLDQFTNI 353

APPLICANT: IRIE, RYOTARO

APPLICANT: TAMECHIKA, ICHIRO

APPLICANT: SEKI, NACHIKO

APPLICANT: YOSHIKAWA, TSUTOMU

APPLICANT: OTSUKA, MOTOTYUKI

APPLICANT: NAGAHARI, KENJI

APPLICANT: MASUHO, YASUHIKO

TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA

FILE REFERENCE: 084335/0160

CURRENT APPLICATION NUMBER: US/10/094,749

CURRENT FILING DATE: 2002-03-12

PRIOR APPLICATION NUMBER: 60/350,435

PRIOR FILING DATE: 2002-01-24

PRIOR APPLICATION NUMBER: JP 2001-328381

PRIOR FILING DATE: 2001-09-14

NUMBER OF SEQ ID NOS: 3381

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 1772

LENGTH: 783

TYPE: PRT

ORGANISM: Homo sapiens

US-10-094-749-1772

Query Match 4.9%; Score 339; DB 4; Length 783;

Best Local Similarity 21.1%; Pred. No. 9.3e-20;

Matches 190; Conservative 161; Mismatches 333; Indels 216; Gaps 41;

Qy 444 LLEQLERHLQVSPQAORNTSLQDICVAPLNPDNTSLYDCCINSLLQYFON----NRT 499

Db 10 ILKHAADVTKIQV--PRPGNFYTFAHICI--LNNDKTCIVDDIVH-VLEELKNARATWRT 64

Qy 500 LLLLTANQTLMGTSQVDKWDHFLYCANAPLT-FKDTALALSCWADYGAVPFPLAIG 558

Db 65 NFAIT-----YPIHLKDGRA-----VYNGHQLGG 89

Qy 559 --YKGD-YSEBALIMTFSLNNYPAGDPRLAQAALWBEAFLEEMRAFQRRMAGFQVTF 615

Db 90 VTVHSKDRVKSAAEQITVYLOSINSLNDWAER--WESSFCDTIVRLFKSNSKVKMYPY 147

Qy 616 TARSLEDEINRTAEDLPIFATSYIVIFLYISLALGYSYSSRVMVDSKATLGLGGVAV 675

Db 148 TSS-SLREDQKTSRVSRYLVTSILV--VTMAILCCS--MQDCVRKPKWGLLGLVT 201

Qy 676 VLGAVMAAGFFSYGIRSLVLTVQVPLVLSVGADNIFIFVLEYQRLPRRGPPEVXH 735

Db 202 ISLATITAGIINLTGKYNSTFLG-VPFVWLGHGLYGTFFEMLSW---RKTREDQHVK 256

Qy 736 IGRALGRVAPSMLLCSLSEA---ICFPLGA--LTPMPAVRTFALTSGLAVIDLFLQMSA 790

Db 257 -ERTAAYVADSMLSFSLTTAMVLTGIGASPTNIEAARIFCCNSCIAIFFNYLYVLSF 315

Qy 791 FVALLSDSKRQASRLDVCCKVP-----QELP-----PPQQGS----- 825

Db 316 YGSSLVFTGYENNYQHSIFCRKVPKEALQEKPAWTRFLLTARFSDTAEGEANTYES 375

Qy 826 GLILGFFQKAYAPFLHWRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLD 885

Db 376 HLLVCFILKRYXCDWITVYKPFVFLVLIYISPALMGYLQVSGSDLSNIVATATQIE 435

Qy 886 YFLFLNRYFEVGAIV--YFVTLTGYNFSS-----EAGNAICSSAGCANNFSTQIKIY 936

Db 436 YTTAQKYFSNYSPIVIGYIESIEYWNYSQVEDVLEYTKGFVIRISFESLYNLRLKLV 495

Qy 937 ATEPPEQSYLAIPASSWDDFDMLTPSSCCRLYISGPNKDKFCPTVNSLNCILKNOWSI 996

Db 496 STGLPKK-----NFTDMLRNS-----FLKAPQSFHQEDLIFS----- 528

Qy 997 TWGSRVRSVEQPHKYPWFLNDRPNIKCPKGLAAYSTVSNLTSDGQVLAASFVAYHKPL 1056

Db 529 -----KKY-----ND-----EVDVVASRMFLVAKTM 549

Qy 1057 K-NSQDYTEALRAARELANANTADUKVPGTDPAPFEPYITVNFYEQYLITLPEGLFM 1115


```
QY 604 QREMA--GMQVTF--ERSLEDEINRTTAEADLPFATSVIVIFLYISLALGSSWSRV 660
DB 354 KNILALKKIEVHVFTSLRQLEPEATSVTV--IPVPHLAYILLILFAVTSCTFRD-----406
QY 661 MVDKATLGLGGVAVVILGAVMAAGPFSYLGIRSSILVILQVPPFLVLSVAGADNIFIVLE 720
DB 407 CIRNKCVAAGFVISAFLAVVSGFGLLHIGV--PFVLIIVANSFPLILGVGVDDMFIMISA 465
QY 721 YQLPRRGPPEVHIGRALGRVAPSMMLCSLSEACFFLGALTPMPAVRTPAITSLGLAV 780
DB 466 WHK--TNLAGDIRE--RMSNVYSKAAVSITITITNIALALYTGIMSSFRSQVCFIYTGTTL 523
QY 781 ILDFLLQMSAFVALLSLSKQASRLDVCCVKPQELP-----PPGQG 824
DB 524 LFCYFYNITCFGAFMALDGKREV-----VCLWLKKAADPKWPFKFCFPPGSGVDEHG 578
QY 825 EGL--LLGPFQKAYAPFLHWTIRGVVLLILFLALFGVLSYMSCHISVGLDQELALPKDSY 882
DB 579 TDIHPSLFRDYFGPFLTRSESKYFVVIYVLYIISIIYGCFFHVQEGDLRLNLASDDSY 638
QY 883 LLDYF 887
DB 639 ITPYF 643

RESULT 73
US-10-890-776A-10
; Sequence 10, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US 10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-890-776A-10

Query Match 4.5%; Score 310.5; DB 5; Length 648;
Best Local Similarity 22.0%; Pred. No. 2.1e-17;
Matches 133; Conservative 107; Mismatches 238; Indels 127; Gaps 24;

QY 334 HT-----LLGQPFQMGVWVSWPLTILVLSVIVVLAAGLVF--TELTDPVLEWSAP 386
DB 115 HTDCLGELLSTRFQWLGWQVGAHPWIFLAPLMLTAALGTGFLYLPKDEEDLEHYTPV 174
QY 387 NSQARSEKAFHDOHFGPFPTNQVILTAPNRSS-----YRDSLLGLGKPFSGIL 436
DB 175 GSPAKAERRFVOGH-----FTTNDYRFSASRRSTRANFVSLVLSVSDSLDLPATFAEVS 230
```

```
QY 437 DLDLLELLELQERLRLHQLQVMSPEAQNRISLQDIC--YAPLNPNTSLYDCCI--NSLLQ 492
DB 231 KLDGAVQDLRV-----AREKGSQIQYQVCARYRAL-----CVPPNPILY 270
QY 493 YQNRNRTALLTANQTLMGQTSQVDWKDHPFLICANAPLTFKQGTALALSCHADYGA--PVF 551
DB 271 AMQVNTKLN-----SSISPPAY-----NHGRHPLY 296
QY 552 PFLAIGY-----KGKDYSEAEALIMTFSLNNYPAGDPRLL--AQAKLWEEAELEEMRAF 603
DB 297 LTGCFPGYILGSLGQQLLRAKANRLIYLYLK--TEDPEYDVQSKQMLTHLDDFTNI 353
QY 604 QREMA--GMQVTF--ERSLEDEINRTTAEADLPFATSVIVIFLYISLALGSSWSRV 660
DB 354 KNILALKKIEVHVFTSLRQLEPEATSVTV--IPVPHLAYILLILFAVTSCTFRD-----406
QY 661 MVDKATLGLGGVAVVILGAVMAAGPFSYLGIRSSILVILQVPPFLVLSVAGADNIFIVLE 720
DB 407 CIRNKCVAAGFVISAFLAVVSGFGLLHIGV--PFVLIIVANSFPLILGVGVDDMFIMISA 465
QY 721 YQLPRRGPPEVHIGRALGRVAPSMMLCSLSEACFFLGALTPMPAVRTPAITSLGLAV 780
DB 466 WHK--TNLAGDIRE--RMSNVYSKAAVSITITITNIALALYTGIMSSFRSQVCFIYTGTTL 523
QY 781 ILDFLLQMSAFVALLSLSKQASRLDVCCVKPQELP-----PPGQG 824
DB 524 LFCYFYNITCFGAFMALDGKREV-----VCLWLKKAADPKWPFKFCFPPGSGVDEHG 578
QY 825 EGL--LLGPFQKAYAPFLHWTIRGVVLLILFLALFGVLSYMSCHISVGLDQELALPKDSY 882
DB 579 TDIHPSLFRDYFGPFLTRSESKYFVVIYVLYIISIIYGCFFHVQEGDLRLNLASDDSY 638
QY 883 LLDYF 887
DB 639 ITPYF 643

RESULT 74
US-10-060-756A-4799
; Sequence 4799, Application US/10060756A
; Publication No. US20030046717A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US 10/060,756A
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4804
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 4799
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-756A-4799

Query Match 4.4%; Score 306.5; DB 4; Length 648;
Best Local Similarity 21.8%; Pred. No. 4.7e-17;
```

Query Match	4.4%;	Score 306.5;	DB 5;	Length 648;
Best Local Similarity	21.8%;	Pred. No. 4.7e-17;		
Matches 132;	Conservative 108;	Mismatches 238;	Indels 127;	Gaps 24;
334 HT-----LLGQFFQCGMTWASWPLRILVLSVIPVVALAAGLVF--TELTTDPVELMSAP	386			
115 HTDCLGLLSRTFQWLGWQVGAHPWIFLLAPLMLTAALGTGFLYLPKDEEDLBEHPTV	174			
387 NSOARSEKAPHDQHPFPFRFQTNVITAPNRSS-----YRDSLLLGKPNFSGIL	436			
175 GSPAKAERRFVQGH-----FTTNDSYRFSASRRSTEANFVSLLVSVSDSLDDPATFAEVS	230			
437 DLDLLLELLELQELRHLQVWSPSAQRNLSLDIC--YAPLNPNTSLSDCCI--NSLLQ	492			
231 KLDGAVQDLRV-----AREKSGQIQVQVCARYAL-----CVPNPILY	270			
493 YFQNNRTLLLTANOTLGMQTSQVDWQKHFLYCANAPLTFKDGDTALALSCMADYGA-PVF	551			
271 AWQNKTKLNL-----SSISFPAY-----NHRGHPLY	296			
552 PFLAIGGY-----KGKDYSEAEALIMTSLNNYPAGDPRKL-AQAKLWEAEFLMERAF	603			
297 LTGFFGGYILGSGMGQQLLRKAKMRLLYLK--TEDPEYDVQSKOWLTHLLDQFTNI	353			
604 QRRWA--GMQVTFETA--ERSLEDEINTKTAEDLPIEATSYIVIFLYISLALGSYSWSKV	660			
354 KNILALKKIEVHPHTSLSRQLEPATSVTV--IIVFHLAYILILFATVSCERPD-----	406			
661 MVDSKATLGLGGVAVLVGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIVLVE	720			
407 CIRNMCMVAAPGVLSAFILVVGSGGLLHIGV-PPVILIVANSPFLILGVGVDMDPIMISA	465			
721 YQRLPRRPGEPREYHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTTALTSGLAV	780			
466 WHK--TNLADDIRE-RMSNVVSKAAVSITITITILALYTGIMSSFRSVQCFCIYTGML	523			
781 ILDFLLQMSAFVALLSLDSKQEAERLDDCCVCKPQELP-----PPGOG	824			
524 LFCYFYNITCFGAFFALDGKEV-----VCLCWLKAKDKWPSFKKCFCPGSGVPDEHG	578			
825 EGL--LLGFFQKAVAPFLLHMTTRGVWLLLFALFLFGVSLYSMCHISVGLDQELALPKDSY	882			

```

RESULT 76
US-09-864-761-34826
; Sequence 34826, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Shaaron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

```


QY 1274 ALBOKRABEA 1283
Db 118 PIEQSQPPSA 127

RESULT 79
US-10-415-934-8
; Sequence 8, Application US/10415934
; Publication No. US20040030099A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN
; FILE REFERENCE: LIO204 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/415,934
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/245,565
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: US 60/245,564
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: US 60/245,572
; PRIOR FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-934-8

Query Match 3.9%; Score 266.5; DB 4; Length 792;
Best Local Similarity 21.1%; Pred. No. 28-13;
Matches 140; Conservative 116; Mismatches 259; Indels 149; Gaps 26;

QY 687 PSYLGIRSLVLQVVPFLVL-SVGADNIFIVLEYQRLPRPGPEVHIGRALGRVAP 745
Db 227 FDFLESVLVIAKAVHNSLSQHGKGVFELSGWR--TKENLPFKORIADAYSDVMV 284

QY 746 SMLLCSLBAICFFIGA--LTPMPAVRTFALSGLAVIDFLQWSAFVALLSLQKQE 803
Db 285 TYTMTSSLVITFGMGASPTNIEAVKVCQNMCSILLNIFYIFSGCLVFAQGLEQ 344

QY 804 ASRLDVCCVKP-----QELPPGGE-----GLLGFQKAYAPL 840
Db 345 NRYHSIFCCIKFSAEYLDKPVWFQVMSDGHQOOSHHTETNYQHFIQHFLREHNEWI 404

QY 841 LEWITRGVLLFLALFGVLSYMSCHISVGLQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 405 TWIYKPFVILYLIYASFSGCLOISDGANIINLLASDSPSVSYAMVQKYFNSYSPV 464

QY 901 ----YFVTTILGY-NFSSEAGMNAICSSAGCNFSFTQKIQYATEFPBQSYLAIPASSWDD 956
Db 465 IGFVYVEPLEYWNSSVQDLRLCS-----GFT-----AVSWVEQ 499

QY 957 FIDWLTSPSCCRLYISGPNKDFCPVNSLNCNCSITMGSV--RPSVQFHKYLPW 1014
Db 500 YQOFLKVS-----VSANKSDP-----ISVLOSSFLUKKPEFQHF----- 534

QY 1015 FLNDRBNIKCPKGLAAYSTSVNLTSQGOVLASRFMAYHKPLKNSQ-DYTEALRAARELA 1073
Db 535 -----RNDIIFSKAG-----DESNIIASRLVIVARTSRDKQKEITEVLEKRLPL- 578

QY 1074 ANITADLRKVPGDPAPEVFPYITNVFYEY-----LTLPEGLFMLSCLVPTTAVSC 1128
Db 579 -SLSKSIRFIV-FNPSF-----VFMDHYSLSVTVPVLIAG-FGVLLVILITLTFPLVI 626

QY 1129 LLIGLRLRGLNLLSIVMLVDTVGFMAIWDISYNVSLINLAVSAGMSVEFVSHITS 1188
Db 627 HPLG-----NFWILSVTSIELVGLMLTNVMDOCISILCLITYITNFAIDHCAPLLFT 681

QY 1189 FAISTKPTMLERAKEATISMGSAVFAGVAMTNLPGIIVLGLAKAQIQLIQLIFFFLNLLITL 1248
Db 682 FVLATEHTRTQCICKSSLQDHGTAL-----LQNTVSFLI-----GLVPLLFPVSNLTFIL 730

QY 1249 L-----GLLHGLVFLPVILSVGPDVNPALALEOKRAE-----EAVAAMVVASCPNH 1295
Db 731 FKCLLLTGGCTLLHCFVILPVLFITFFPP--SKKHHKKKRAKREKEIEICIEIQENPDH 788

QY 1296 PSRV 1299
Db 789 VTTV 792

RESULT 80
US-10-424-599-181755
; Sequence 181755, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181755
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)-(229)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135138C.1.pep
US-10-424-599-181755

Query Match 3.8%; Score 264.5; DB 4; Length 229;
Best Local Similarity 28.4%; Pred. No. 4.3e-14;
Matches 75; Conservative 44; Mismatches 102; Indels 43; Gaps 8;

QY 369 GLVFTLTTDPVELWSAPNSQARSEKAFHDQHGFPTNQTQVILTPAPNRSYRYSLLLG 428
Db 4 GLLREFAEETREPKLVWPGSKAAEKEFFDLSHAFYRIEQILITATIPESKHGKPSIIT 63

QY 429 PKNFGIILDLLELLELQERLHLQVWSPAQRNLSLODICYAPLNPDTSLYDCCIN 488
Db 64 EEN-----IELFEIQEKVDGIR-----ANYSGLLVSLSDICLPLGD-----DCASQ 106

QY 489 SILQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTFKDGDTALALSCMADYGA 548
Db 107 SILQYFQ-----MDPDNYDLRRCXHAERYWFQHYTSTE-TCFSAPKA 147

QY 549 PVFPFLAIGGYKGDYSEAEALIMTFSLNN--YPAGDPR---LAQAKLWEEAFLEEMRAP 603
Db 148 PLEPTTALKXSGNNYSEASAFVITVPVNNAITKVGDENGKAIXGKAFIQLAKEBLLPM 207

QY 604 QRRMAGMFQVTTAERSLEDEINR 627
Db 208 VQ--SNNLTSPSTESSIEBELKR 229

RESULT 81
US-10-425-115-201210
; Sequence 201210, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201210
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(267)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_11508C.1.pcp
US-10-425-115-201210

Query Match 3.7%; Score 259; DB 4; Length 267;
Best Local Similarity 29.5%; Pred. No. 1.6e-13;
Matches 77; Conservative 38; Mismatches 108; Indels 38; Gaps 13;

QY 20 QSEPYTHIQPGYCAFYDECGKPELSGLMTLSVSCLSNTPARKITGDHLI--LLOKI 77
DB 34 QNSGSRVPAEGYCSMYGICQARSD--GKVLNCANAT-----KAVKPDTLFSSRIQSL 84
QY 78 CPELYTGPNQACCSAKQLVLSLASLITKALLTRCPACSDNFVNLHCHNTCSPNOSLFI 137
DB 85 CPTI--TG--DVCCTVDQFDTLHQVQVAVPFLVGCPCACLRNLFNFCMSCSPNOSLFI 140
QY 138 NVTRVQLGAGQLPAAVVAEAFYOHGFAEQSYDCSRVRPAAATLAVGTMCVGVGSAIC 197
DB 141 NTVSVKQINSTM--TWGDIDYVYTHYGBELVNSCKDVKFGTINTAMPFI-----GAGAK 194
QY 198 NAQRWLNFP---QSDTNGNGLAPLDTITFHLLEPGQ--AVGSGIQPLNGBVARCNESQDDVA 252
DB 195 TYKDWLAFIXRQANLNKPGSPYLITE-----GSDPDSGSKVFLXSTIYSC-----GTPFL 245
QY 253 TCS-CODCAAS--CPAIRPQ 270
DB 246 XCSXCGDCPSSXKAGSLLPQ 266

RESULT 82
US-10-408-765A-1316
; Sequence 1316, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1316
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1316

Query Match 3.6%; Score 252; DB 4; Length 505;
Best Local Similarity 20.9%; Pred. No. 1.8e-12;
Matches 124; Conservative 105; Mismatches 219; Indels 146; Gaps 24;

QY 756 ICFELGA--LTPMPAVRTALTSLGLAVILDFLLQMSAFVALLSLDSKQESRLDVCCV 813
DB 8 ITFGMGASPTTIEAVKVCQNMVCSILLNYPIFSFGSLVFAGQLEQNRYHSIFCK 67

QY 814 KP-----QELPPPGGGE-----GLLGGFFQKAYAPFLLHWTIRGWL 850
DB 68 IPSAEYLDLRKPVWFQTVMSDGHQTSHTHETNPQHHPHQHFLREHYNEWITNIYVKFPV 127
QY 851 LFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV---YFVTTLG 907
DB 128 ILYLIYASFSPMGCLQISDGANIINLLASDPSVSAMVQOKYFSNYSPIGVFYVPEPLE 187
QY 908 Y-NFSEAGMNAICSSAGCANNPSPFTQKIQVATFEPPQSYLELAIPASSWVDPDFDLWLTSSC 966
DB 188 YNNSVQDDLRRLCS-----GFT-----AVSWVEQYQFLKVSN- 221
QY 967 CRLYISGPNKDFCPSVTNLSNLCNKMCSITMGSV--RPSVEQFHKYLPWFLNDRPNIKC 1024
DB 222 ----VSANKKSDP-----ISVLQSSFLKKPFGHF-----RNDIIF 253
QY 1025 PKGGLAAYSTSVNLTSDGQVLAASFMYHKLKNSQ--DYTEALRAARELANITADLRKV 1083
DB 254 SKAG-----DESNIIASRLYLVA RTSRDKQKEITEVLEKRLPL--SLSKSRPI 300
QY 1084 PGTDPAFEPVPPYITNVFYEQY-----LTLPEGLFMLSICLVPTTAVSCLLGLDLRS 1138
DB 301 V-FNPSF-----VFMDHYSLSVTVPVLIAG-FGVLLVLITLFTFLVIHPLG-----N 344
QY 1139 LMLLSIVMLVDTVGMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWL 1198
DB 345 FWLILSVTSIELGVGLMTLVNVDMDICISILCIYITLNFDAIDHCAPLFTFVLATEHRT 404
QY 1199 ERKAEATISMGSAFAGVAMTNLPGLVLGLAKAQIQLIPFFRLNLLITLL----- 1249
DB 405 QCIKSSLDQHGTAI-----LQNVTSPLI-----GLVPLLFPVSNLTFTLFCCLLTGGC 453
QY 1250 GLLHGLVLPVILSVYGVDPNPALEQKRAE-----EAVAAMVAVSCPNNHPSRV 1299
DB 454 TLLHCFVILFVLTTPPP--SKHHKKKKRKRKRERIECIEIQENPDHVTTV 505

RESULT 83
US-10-432-613-2
; Sequence 2, Application US/10432613
; Publication No. US20040048282A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN
; FILE REFERENCE: L10234 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/432,613
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: US 60/251,032
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-432-613-2

Query Match 3.6%; Score 245.5; DB 4; Length 1491;
Best Local Similarity 20.6%; Pred. No. 3.6e-11;
Matches 247; Conservative 150; Mismatches 494; Indels 307; Gaps 52;

QY 213 LAPLDITP-HLLEPGQAVSGIQPLNEGVARCNESQGDVATCSQDCAASCFAIARPOA 271
DB 179 LAPAHFTYPRALQEQYQG-GSSLPGI-----GDRAALCSHGSSLSPSGA---PSQ 223
QY 272 LDSTF-----YLGQMPGS-----LVLIILCSVFAVTVLIVGFR 306
DB 224 RDGTWKPPAVQHHVSVRQERAFQMPKSVSQLIAEMPVAVLMCLAVIFLCTLAGLGR 283
QY 307 VAPARDKMWDPKKGTSLSKLSFSTHTLLGQPFQGMGTWVASWPLT-----ILVLSVIP 362
DB 284 L-PDFSKPLLGPEPRDITDIGSKLV-----VVRALQALTGPRLKLLPLS--P 325

[illegible]

```

; GENERAL INFORMATION
;
; APPLICANT: Bayer AG
; TITLE OF INVENTION:
; FILE NUMBER: L100
; CURRENT APPLICATION:
; CURRENT FILING DATE:
; PRIOR APPLICATION N:
; PRIOR FILING DATE:
; PRIOR APPLICATION N:
; PRIOR FILING DATE:
; PRIOR APPLICATION N:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS:
; SOFTWARE: PatentIn
; SEQ ID NO 2
; LENGTH: 785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-934-2

```

Query Match	3.3%	Score 230;	DB 4;	Length 785;
Best Local Similarity	18.6%;	Pred. No. 3e-10;		
Matches	155;	Conservative 110;	Mismatches 273;	Indels 294;
Gaps	32;			
Qy	334	HT-----LLGQFFQGGTGWASWPLTILVLVSIPVVVLAAGLVF--	-TETTTDPVLEWSAP	386
Db	115	HTDCEGLSRTQWLQWQGAIPWFLAPLMLTAALGTGFLYLPKDBEEDLEEHYTV	174	
Qy	387	NSQAREKAFHDQHFGPFRTNQVILLTAPNRSS-----YRDSLLLLGPKNPSGIL	436	
Db	175	GSPAKAERRFVQGH-----FTTNDSYRPSASRRSTEANFVSLLVVVSYSDSLLDPATFAEVS	230	
Qy	437	DLDLLLELLEQLERHLQVWSPEAQNISLQDIC--YAPLNPDNTSLYDCCI--NSLLQ	492	
Db	231	KLDAVGQDLRV-----AREKSGIQIYQVQCARAL-----CVPNPILY	270	
Qy	493	YFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGDTALALSCMADYGA--PVF	551	
Db	271	AWQVNTKLNL-----SSISFPAY-----NHGRHPLY	296	
Qy	552	PFLAIGY-----KQKQYSEAEALIMTSLNNYPAGDPRL--AQAKLWEAFLEEMRAP	603	
Db	297	LTGFFGVIYLLGSLGMLLRKAKEMLLYLK---TEDPEYDVQSKOMLTHLLDQDTNI	353	
Qy	604	QRMA-----GMFQVTTAERSLEDEINRTTAEDLPFATPSYIVIFLYISL	649	
Db	354	KNILALKKIEVPGVGIGQGEKRVHTLSRQLEFATSVTVIPVFLAYILLIPAVT	413	
Qy	650	ALGSYSWSRNVVDSKATLGLGGVAVVLGAVMAAGPFYSYLGIRSSILVILQVPLVLVSU	709	
Db	414	SCRFD-----CIRKNKVAAGFVISAFLAVVSFGLLLHIGV--PFVIVANSFPLILGV	467	
Qy	710	GADNIFIPVLEBYQLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFIAGALTPMPAV	769	
Db	468	GVDDMFIMISAWHK--TNLAGDIRE--RMSNVYSKAAVSITITITINILALYGI	518	
Qy	770	RTFALTSLGLVILDFLLQMSAFVALLSLDSKRQEASRLDVCCVKPQLPFPFGQEGILL	829	
Db	519	MSF-----MSF-----	522	
Qy	830	GPFOKAYAPFLHMITRGVILLFLALFGVLSYSWCHISVGLDQELALPKDSYLLDYFLF	889	
Db	523	SIYGFHVQEGIDLRLNLSDDSYITITPYFNV	552	
Qy	890	LMRYF--EVGAPVYFVTTLGYNPSSEAGNNAICSSAGCNPFSTQKIQVATEPEOSYLAI	948	
Db	553	BEYFSDYGPVMVITVKVDYWDK-----DVRQKLECNCTKIFKNVYVD	597	
Qy	949	P--ASSWDDFDID-----WLTPSSC-----CRLYISGP--NKDKFCP	981	
Db	598	KNLTFELWDAYVQYLKGFPMNIVWEKLSNCNVAINQTLWLVKANASIPLYGPLNNKMRGP	657	
Qy	982	S-----TVNSLNCIKNC-----MSITMGSVRPSVEQPHKLPWFLDRPNIK	102	

Db 658 GGIVHTRILVERLTCLLTSVLAVNLKVALWLSISPPQLRPSSELL-----TQ 704
QY 1024 CPGGGLAAYS-----TSVNLTSQGVLASRFMAYHKPLKNSQDYTE 1064
Db 705 GPRGTLAHLSQKSLSPFESFANFCLTADSKARSQALL---PSGNCPIYAE 753
RESULT 85
US-10-017-161-2024
; Sequence 2024, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2024
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-2024

Query Match 3.3%; Score 225; DB 4; Length 542;
Best Local Similarity 20.4%; Pred. No. 4.6e-10;
Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;
QY 744 APSMLCSLSEA---ICFPLGA--LTPMPAVRTFALTSGLAVIDLDFLLQMSAPVALLSLD 798
Db 23 AUSMLSFSLTAMLYVTFGIGASPTNIEARIFCCNSCIAIFFNLYLVLSFGSSLVFT 82
QY 799 SKRQASRLDVCCVKP-----QELP-----PPQGE-----GLLGGFFQ 833
Db 83 GVIENNYQHSIFCRKVPKEALQEKPAWYRFLLTARFSDTAGEEANTYVESHLLVCFLK 142
QY 834 KAYAPFLHWTGRVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRY 893
Db 143 RYCDWITNTYVKPFVFLYIISPALMGYLQVSESDLSNIVATATQIETTTAAQKY 202
QY 894 FEVGAPVYFVTTLGYNFSEAGMNAICSSAGCNSFTQKIQVATFPPQSYLAI PAS-- 951
Db 203 FSNYSPVI-----GFYIESIEYWNITSVQEDVLETKGFV 237
QY 952 --SWDDDFIDMLTPSSCCRLYIS--GPNKDKFCPSTVNSLNCNKNCMSITMGSVRPSVEQF 1008
Db 238 RISWPESYLNYLR-----KLVSTGLPKNFTDMLRNSP--LK-----273
QY 1009 HKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLK--NSQDYTEALR 1067
Db 274 ---APQSFHQBDLIFSK-----KYNDEVD-----VVASRMFLVAKTMTNREELYDLLE 320
QY 1068 AARELAANTADLRKVPGTDPAFEPVPTITVTFVEQYLTILPEGLFMLSCLVPTFAV- 1126
Db 321 TLRL--SVTSKVKFIV-FNPSF-----VYMDRYASSL--CAPLHNSCISALFLIF 366
QY 1127 -SCLLGLDLRLSGLNL---LSIVMLVDTVGFMALWDISYNVAVSLINLVSAGMSVRFV 1182
Db 367 FSAFLVA---DSLINVMTLTVSVFVGIGFTLWKVELDCISVLCLYIYINITYDNC 422
QY 1183 SHITRSFAIS---TKPTWLERAKRATISGSAVFAGVAMTN-----LPGILVLGLAKAQL 1234
Db 423 APMLSTFVLGKDFTRTKWKVNALEV-----HGVALQSYLCVIVGLIPLAAVPSNL 473
QY 1235 IQIFFRNLMLTLGLLHGLVFLPVILSYVGPVDNVPALAEQKRAE 1282
Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFPLPS-----KKRKEK 513

Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFPLPS-----KKRKEK 513
RESULT 86
US-10-292-798-1670
; Sequence 1670, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1670
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1670

Query Match 3.3%; Score 225; DB 4; Length 542;
Best Local Similarity 20.4%; Pred. No. 4.6e-10;
Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;
QY 744 APSMLCSLSEA---ICFPLGA--LTPMPAVRTFALTSGLAVIDLDFLLQMSAPVALLSLD 798
Db 23 AUSMLSFSLTAMLYVTFGIGASPTNIEARIFCCNSCIAIFFNLYLVLSFGSSLVFT 82
QY 799 SKRQASRLDVCCVKP-----QELP-----PPQGE-----GLLGGFFQ 833
Db 83 GVIENNYQHSIFCRKVPKEALQEKPAWYRFLLTARFSDTAGEEANTYVESHLLVCFLK 142
QY 834 KAYAPFLHWTGRVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRY 893
Db 143 RYCDWITNTYVKPFVFLYIISPALMGYLQVSESDLSNIVATATQIETTTAAQKY 202
QY 894 FEVGAPVYFVTTLGYNFSEAGMNAICSSAGCNSFTQKIQVATFPPQSYLAI PAS-- 951
Db 203 FSNYSPVI-----GFYIESIEYWNITSVQEDVLETKGFV 237
QY 952 --SWDDDFIDMLTPSSCCRLYIS--GPNKDKFCPSTVNSLNCNKNCMSITMGSVRPSVEQF 1008
Db 238 RISWPESYLNYLR-----KLVSTGLPKNFTDMLRNSP--LK-----273
QY 1009 HKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLK--NSQDYTEALR 1067
Db 274 ---APQSFHQBDLIFSK-----KYNDEVD-----VVASRMFLVAKTMTNREELYDLLE 320
QY 1068 AARELAANTADLRKVPGTDPAFEPVPTITVTFVEQYLTILPEGLFMLSCLVPTFAV- 1126
Db 321 TLRL--SVTSKVKFIV-FNPSF-----VYMDRYASSL--CAPLHNSCISALFLIF 366
QY 1127 -SCLLGLDLRLSGLNL---LSIVMLVDTVGFMALWDISYNVAVSLINLVSAGMSVRFV 1182
Db 367 FSAFLVA---DSLINVMTLTVSVFVGIGFTLWKVELDCISVLCLYIYINITYDNC 422
QY 1183 SHITRSFAIS---TKPTWLERAKRATISGSAVFAGVAMTN-----LPGILVLGLAKAQL 1234
Db 423 APMLSTFVLGKDFTRTKWKVNALEV-----HGVALQSYLCVIVGLIPLAAVPSNL 473
QY 1235 IQIFFRNLMLTLGLLHGLVFLPVILSYVGPVDNVPALAEQKRAE 1282
Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFPLPS-----KKRKEK 513

RESULT 87									
US-10-415-934-14									
; Sequence 14, Application US/10415934									
; Publication No. US20040030099A1									
; GENERAL INFORMATION:									
; APPLICANT: Bayer AG									
; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN									
; FILE REFERENCE: LIO204 Foreign Countries									
; CURRENT APPLICATION NUMBER: US/10/415.934									
; CURRENT FILING DATE: 2003-05-05									
; PRIOR APPLICATION NUMBER: US 60/245,565									
; PRIOR FILING DATE: 2000-11-06									
; PRIOR APPLICATION NUMBER: US 60/245,564									
; PRIOR FILING DATE: 2000-11-06									
; PRIOR APPLICATION NUMBER: US 60/245,572									
; PRIOR FILING DATE: 2000-11-06									
; NUMBER OF SEQ ID NOS: 16									
; SOFTWARE: Patent in version 3.1									
; SEQ ID NO 14									
; LENGTH: 563									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-415-934-14									
Query Match 3.3%; Score 225; DB 4; Length 563;									
Best Local Similarity 20.4%; Pred. No. 4.9e-10;									
Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;									
QY	744	APSMLLCSLSEA---ICPFLGA---LTPMPAVRTEALTSGVLVILDFLLQMSAFVALLSLD	798						
DB	44	ADSMLSFTTAMYLVTFGIGASPTNIEAARIFCCNSCIAIFNNLYLVLSFGYSLVFT	103						
QY	799	SKRQESRLDVCCKVP-----QELP-----PPQGE-----GLLGGFQ	833						
DB	104	GNIENYQHSIFCRKVPKEALQEPAYRFLLTARFSDTAEGEEANTYESHLLVCPLK	163						
QY	834	KAYAPFLHWITRGVLLFLALFGVLSYSCMHISVGLDQELALPKDSVLLDYFLNRY	893						
DB	164	RYVCDWITTVYKPFVVLVLIYISPALMGYLQVSGSDLSNIVATQTITETTAQKY	223						
QY	894	FEVGAVPYFTTLTGYNFSEAGNACSSAGCNNEFSFTKIOVATEFPQSLAIPAS--	951						
DB	224	FSNYSPI-----GFYIESIEVWNTSVQEDVLEYTKGFV	258						
QY	952	--SWDDFDLWLPSSCCRLIYIS-GPNKDKFCPSTVNSLNCNKMSTITWGSVRPSVEQF	1008						
DB	259	RISWFSYLNLYR-----KLVNSTGLPKKFTDMLNSP--LK-----	294						
QY	1009	HKYLPWFLNDRPNIKCPKGGLAAYSTSVNLTSDGQVLASRFMAYHKPLK-NSQDYTEALR	1067						
DB	295	---APQFSHFQEDIIFSK-----KYNDEVD-----VVASRMFLVAKTMTNREELYDLLE	341						
QY	1068	AARELANITADLRKVPCTDPAFEVPPYITITVFEQYLTILPEGLFMLSCLCVPTFAV--	1126						
DB	342	TLRRL--SVTSKVKFIV-FNPSP-----VMDRYASSL--GAPLHNSCISALFLFP	387						
QY	1127	-SCLLGLDLSRLNLN---LSIVMILVDTVGFMALWDISYNAVSLINLVSAGVMSVEFV	1182						
DB	388	PSAFLVA---DSLINWITLVVSVEFGVIGFWTLMKVVELDCISVLIIYGINYTIENC	443						
QY	1183	SHITRSPAIS---TKPTWLRERAKATISMGSAVPAGVAMTN-----LPGILVLGLAKAQL	1234						
DB	444	APMLSTFVLGKDFTRTKVWKNALFV-----HGVAILQSVLCYIVGLIPLAAPSNL	494						
QY	1235	IQIPFFRLNLLITLLGLLGLVFLPVILSVYVGPDPNPALEOKRAE	1282						
DB	495	-TCTLFCFLCIAFVTFHCFAPILPVILTLFPPS-----KKGKKEK	534						
RESULT 88									
US-10-432-613-3									
; Sequence 3, Application US/10432613									
; Publication No. US20040048282A1									
; GENERAL INFORMATION:									
; APPLICANT: Bayer AG									
; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN									
; FILE REFERENCE: LIO234 Foreign Countries									
; CURRENT APPLICATION NUMBER: US/10/432.613									
; CURRENT FILING DATE: 2003-06-04									
; PRIOR APPLICATION NUMBER: US 60/251,032									
; PRIOR FILING DATE: 2000-12-05									
; NUMBER OF SEQ ID NOS: 7									
; SOFTWARE: Patent in version 3.1									
; SEQ ID NO 3									
; LENGTH: 1218									
; TYPE: PRT									
; ORGANISM: Drosophila melanogaster									
US-10-432-613-3									
Query Match 3.1%; Score 214.5; DB 4; Length 1218;									
Best Local Similarity 18.7%; Pred. No. 1.3e-08;									
Matches 210; Conservative 185; Mismatches 412; Indels 317; Gaps 50;									
QY	374	ELTDPVELWSAPNSQARSEKAFHQHGFPPFRTNOVILTAPNRSSRYRYDSLILGPKNFS	433						
DB	222	QAATLPTDGMADMRROPIEGFCD-----SSPRKEYSHFVVQRIGPNATD	267						
QY	434	GILDLDLLELLELQERLRLHLQVMSPEAQRNISLQDIC-----YAPLNPDNTSLYDCC	486						
DB	268	SLFDLNGLLAMCQLQDQITTEVPSYRAFCEPEMLTTECCRPWSLPNPAAMLANKSSCFD--	325						
QY	487	INSLQYFQNNRLLLTANQTLMGQTSQVDWKHF-----LYCANAPLTFKDCGTALAL	540						
DB	326	-----LTATEDVTSLHTLLGCEYFHDLMKMDNHCNEIP-----	358						
QY	541	SCMADYGAPVFPFLAIGYKGYKQVGEAREALIMTFLSNVYPAGDPRLAQAKLWEAEFLBEM	600						
DB	359	HCRAP-----BECKRLNIVFNLPF-----LTDPSFIKSN	388						
QY	601	R-----AFORMMAGMQVTFTAERSLEDEINRTTABDL-----	633						
DB	389	DSNVYLKYAMIFIPVAQSNRLPLPHEWEDVE--LINELVEVVAMDGLGLENELFNELLIT	446						
QY	634	PIFATSYIVIFLYISIALGSSYSNRVMDSKATIGLGGVAVVLGAVMAAMGFESVLGIR	693						
DB	447	DVWLVSGLGTFYMASVWLYTGSAPITLMSCAVICSFL--GLAYFFVAIVLEFEFFPYMULL	505						
QY	694	SSLVILQVVPFLVLSVGDNIPIFVLEY-----QRLPRR-----PGE	730						
DB	506	AVVVI-----IGIGADDDVFLFKIWHCVLTERFSNRCITLTQSQSALPTLENSDHT	557						
QY	731	PREVHIGRALGRVAPSMLLCSLSEAICFFLGLALTMPAVRTFALTSGSLAVILDFLLQMSA	790						
DB	558	SLENIMALTMRHAAASMFVTSLTTAGAFYASYSSTAIKCFGIFAGTVVVTNYLLMITW	617						
QY	791	FVALLSLDSKQESRLDVCCKVQELPPPGGEGLLLGPFQKA---YAPFLHWTTRG	847						
DB	618	LPASVSI-MERLFATRMS---CHHPMSIK-----LIHACKKSINRQCMFEBCITKS	665						
QY	848	VV---LLLFLALFGVLSYSCMHISVGLDQELALPKDSVLLDYFLNRY-PEVGAPV----	900						
DB	666	IMNYAVLWLLIFG-ALGASAVIVFWYPGLOLEKSH---FQLFVSKHPFEVYSLLKQOF	721						
QY	901	YFVTTL-GY-NFSSEA-----GMNAICSSAGCNNEFSFTKIOVATEPQEQSLAIPASSWV	954						
DB	722	WFEKPLQAYENPKWMHFMVWGVAVDGDYTNPSYGH-LHYDNNEFNVSSR---PAQLWI	777						
QY	955	DDFIDMLTPSSCCRLIYISGNKDKFCPESTVNSI--NCL-----KNCMSITMGSVR-	1002						
DB	778	LDF-----CQSVRQRP-----FYKETLGLMLLPNCFIENLDYMKRCRID-DMDSTRK	823						
QY	1003	-----PSVEQFHKLFPWFLND-----RPNIKCPR-----	1026						
DB	824	DRSPCCDAQPFPEPHI--FEYCLPQSISSNMVDTFFRPGVAGPKFASAPRLETEDYLGMS	881						
QY	1027	-GGLAAYSTSVNLTS-DGQVLASRF---MAYHKPLKNSQDYTEALR-----AAREL	1072						

Db 882 GNEA5YSTNGSTPLLLKALVIEFESNVAISTYIYANIRQFYSEVHEWFMQMLKTAPEL 941
Qy 1073 AAN-ITADLRKVPGTDPFAFEVFPYITITNVFYEQYLTLPEGLPMLSLCLVPTFAVSCLLL 1131
Db 942 QCGWFTSDLK-----FYNVQDTLSHDTFVAICLAASAIAV-LLCF 981
Qy 1132 GLDLRSGLLNLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMVSFVSHITRSPAI 1191
Db 982 TVNILISIVAVLTVSLISIFNTVAVLILGLWQLNILESIAVSTAIGLAVDFSLHYGIHYRM 1041
Qy 1192 STKPTWLERAKEATIS--MGSAPVAGVAMTNLPGILVGLAKAQLIQIFPERLNLITLL 1249
Db 1042 SPVKERL--AATQFVLSRIIGPTVMAATT--TGLAGGIMMASNILPYIQIGVFL--VVVMIV 1097
Qy 1250 GLLHGLVFLPVILSYVGPD-----VNPALALEQKRAEAV 1284
Db 1098 SMFYATFLLMSLLRVAGPQHGFLKWLKWSKSSGSKFYERKPSQVIASEQLLTPSS 1157
Qy 1285 AAVWASCPNHPRSVSTADNIYVNHSPFGSIKAGAIINFLPNN 1328
Db 1158 AIVELANSETHELESINSNLIKTISS--GIESAHALSS-LPRD 1197

RESULT 89

US-11-097-143-3123
; Sequence 3123, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3123
; LENGTH: 1218
; TYPE: PR1
; ORGANISM: DROSOPHILA
US-11-097-143-3123

Query Match 3.1%; Score 214.5; DB 6; Length 1218;
Best Local Similarity 18.7%; Pred. No. 1.3e-08;
Matches 210; Conservative 185; Mismatches 412; Indels 317; Gaps 50;

Qy 374 ELTDPVELWAPNQAQSEKAFHQHGFPPRTNQVILTAPNRSRYDSLLGPKNFS 433
Db 222 CAATLFTDGMHRRQPTIEGFFCD-----SSPRKEYSHFVVQRIQGNATD 267
Qy 434 GILDLDLLELELOERHLQVWSPQAQNSLQDIC-----YAPLPNDNTSLYDCC 486
Db 268 SLFDLNGLLAMCOLQDQITEVPSYAFCEPEMLTTECCRPWSLPNYAALANKSCFD-- 325

RESULT 90

US-10-168-428-2
; Sequence 2, Application US/10168428
; Publication No. US20030165897A1

Qy 487 INSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHF-----LYCANAPLTPKDGATLAL 540
Db 326 -----LTDEDVTSHTLLLGCVYFPHDKMDNHCNEIP----- 358
Qy 541 SCHADYGAVPFPLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLWEEAFLEEM 600
Db 359 HCRAP-----BECKRLNIVFVNLN-----LTDPSFIKSN 388
Qy 601 R-----AFORMAGMFQVTFATRSLEDEINRTTAEOL----- 633
Db 389 DSNVLYKIAMIFIPVAQSNELLPLFHEWEDVE--LINELVEVVAMDGLLENELFNEILLT 446
Qy 634 PIPATSYIVIFLYISIALGSYSYSGRWVSDSKATGLGGVAVVLGAVMAAMGPFYSLGR 693
Db 447 DVMVLSLGGTFVMA5WLVYTGSAFTLMSCAVCFSL-GLAYFFYAIIVLEPEFFPYMNL 505
Qy 694 SSVILQVVPFVLVLSVGDNIPIFVLEY-----ORLPRR-----PGH 730
Db 506 AVVVI-----IGIGADDDVFLKIMHCVLTERPFSNRECTLTQSQSALPTLENSDHT 557
Qy 731 PREVHIGRALGRVAPSMILCSLSEAIICFFPLGALTMPAVRTALTSLGLAVILDFLLQMSA 790
Db 558 SLENIMALTMRHAAASMFVTSITAGAFYASVSSSITAIKCGIFAGTVVTVNYLMTW 617
Qy 791 FVALLSLDSKROEASRLDVCCVPPQELPPQGBGLLGFQKA---YAPFLHWHITRG 847
Db 618 LPASVSI-MERLFAITMS--CHPMSIK-----LIHACKSINRCQMPFECITKS 665
Qy 848 VV-LLFLALRGVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRY-FEVGAPV----- 900
Db 666 IMNYAYLWLLIFG-ALGASSAVIVFVYFGLQLPKESH---FOLFSVKHPFEVYSSLKQOF 721
Qy 901 YPVVTLT-GV-NFSSSEA---GMNAICSSAGCNPFSFTQKIQVATPEPESYLAIPASSV 954
Db 722 WFEKPLQAYENFKMMHFWGVQVDDGDYTNPNSTGH-LHYDNNPNVSSR---PAQLMI 777
Qy 955 DDFIDWLTSPSSCCRLYISGPNKDKFCPTVNSL--NCL-----KNCMSITMGSVR- 1002
Db 778 LDF-----CQSVRQOP---FVKETLGLMFLNCPFIENLDYMKRCID-DMDSTRK 823
Qy 1003 -----PSVEQFHKYLPWFLND-----RPNIKCPK----- 1026
Db 824 DRSPCCDAQPPPEPHI--FEYCLPOSISNMVYDTTFREGVAGPKFAEAPRLETEDYLGMS 881
Qy 1027 -GGLAAYSTSVNLTS-DGOVLASRE--MAYHKPLKNSQDYTEALR-----AAREL 1072
Db 882 GNEA5YSTNGSTPLLLKALVIEFESNVAISTYIYANIRQFYSEVHEWFMQMLKTAPEL 941
Qy 1073 AAN-ITADLRKVPGTDPFAFEVFPYITITNVFYEQYLTLPEGLPMLSLCLVPTFAVSCLLL 1131
Db 942 QCGWFTSDLK-----FYNVQDTLSHDTFVAICLAASAIAV-LLCF 981
Qy 1132 GLDLRSGLLNLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMVSFVSHITRSPAI 1191
Db 982 TVNILISIVAVLTVSLISIFNTVAVLILGLWQLNILESIAVSTAIGLAVDFSLHYGIHYRM 1041
Qy 1192 STKPTWLERAKEATIS--MGSAPVAGVAMTNLPGILVGLAKAQLIQIFPERLNLITLL 1249
Db 1042 SPVKERL--AATQFVLSRIIGPTVMAATT--TGLAGGIMMASNILPYIQIGVFL--VVVMIV 1097
Qy 1250 GLLHGLVFLPVILSYVGPD-----VNPALALEQKRAEAV 1284
Db 1098 SMFYATFLLMSLLRVAGPQHGFLKWLKWSKSSGSKFYERKPSQVIASEQLLTPSS 1157
Qy 1285 AAVWASCPNHPRSVSTADNIYVNHSPFGSIKAGAIINFLPNN 1328
Db 1158 AIVELANSETHELESINSNLIKTISS--GIESAHALSS-LPRD 1197

```

; GENERAL INFORMATION:
; APPLICANT: Universit,t Zrich
; TITLE OF INVENTION: Dispatched Polypeptides
; FILE REFERENCE: D. melanogaster dispatched sequence
; CURRENT APPLICATION NUMBER: US/10/168,428
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 09/468,237
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-168-428-2

Query Match          3.1%; Score 212.5; DB 4; Length 1218;
Best Local Similarity 18.4%; Pred.No. 2e+08;
Matches 208; Conservative 184; Mismatches 409; Indels 329; Gaps 48;

QY   374 ELTDTPEVLSAPNSQARSEKAFHQHGFPPFRTNQVILTPAPNRSSRYSDSLGPKNFS 433
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   222 QAAATLPTDGMWHRQPTEGFFCD-----SSPRKEYSHFVVQRIGPNATD 267

QY   434 GIIDLDDLLLELQLERHLQVWPSPQAQRNISLQDIC-----YAPLPNDNTSLYDC 486
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   268 SLFDLNGALLAMCOLQDITEVPSYRAFCEPMLTTTECCRPMSLNPYAAMLANKSCSDF-- 325

QY   487 INSLLQYFNRRTLTLLLTANQTLMGTOSVDWKHF-----LYCANAPTTFKDGTALAL 540
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   326 -----LTTEDVTSLHTLLGCYEYPHDLKMDNHCHNEIP----- 358

QY   541 SCMADYGAVFPFLAIGGYKGDKYSBAELIMTFPLNNYPAGDPRLPAQAKIWEBAFLSEM 600
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   359 HCRAP-----BECKRLNVFNVLNF-----LTDFSFKSN 388

QY   601 R-----AFORMAGMFQVTTAERSLEDEINRTAEDL----- 633
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   389 DSNVYLKYAMIFIPVAQSNRLLPLPHEWEDVE--LINELVEVMAMDGLGNELNELLLT 446

QY   634 PIPATSYIVFLVYISLAGSYSSWSRWVDSKATLGCGVAVLVGVAVMAAMGFFSYLGIR 693
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   447 DVWLVSJGGTFVWASVWMLYTGSAFILMSCVALCSFL-GLAYFPFAIVLEFEFPFYNNLL 505

QY   694 SSLVILQVPFLVLSVGADNIFIVLEY-----QRLPRR-----PGE 730
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   506 AVVI-----IGIGADDVFLEKIWHCVLTERESNRCTLTTSQSALPTLENSDSTE 557

QY   731 PREVHGIRGVRAPSMLLCSLEACFFLGALTMPAVRTALTSGLAVIDLFLLQMSA 790
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   558 SLENIMALTRHAASMFVTSITAGAFASYSSSITAICFCGIFAGTVVTVNYLLMITW 617

QY   791 FVALLSDSKROEASRLDVCCVKPQBLPPPGQEGELLGFOKA---YAPPELLHWITRG 847
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   618 LPASVSI-MERLFATRMS---CHIPMSIK-----LIHACKSINRCOMPFECSIKS 665

QY   848 VV--LLLFLALFGVSLXSMCHISVGLDQELALPKDSYLLDYFLFLNRY-FEVGAPVYFVT 904
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   666 IMNYAYLWLLIFG-ALGASSAVIVFWPYGLQLEKSH---FOLFVSKHPFEV-----YS 715

QY   905 TLGVNFSSEA-----GNAIALCSSAGCNPFSTFKTKIQTAYATEPEQSYLAI 948
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   716 SLKKQCFWFKEKPQWQAYENPKMHMFVWGQAVDDGDYNPNNSYGH-LHYDNFNVSGR--- 771

QY   949 PASSWVDDFIDWLFPSSCCRILYISGPNKDKFCPSTVNSL--NCL-----KNCMSIT 997
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   772 PAQLWILD-----CQSVROOP-----FYKETLGNLLPNCFIENLDYMKRCID-D 817

QY   998 MGSVR-----PSVEQPHKYLPMFLND-----RPNIKCPC----- 1026
      :|::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB   818 MDSTRKDRSPCCDAQFPPEPHI--FEYCLPOSIGNMYDDTTTTFRPGVAGPKFAEPRALETE 875

QY   1027 -----GGLAAYSITSVNLT-SDQVLASRF---MATHKPLKNSQDYTEALR----- 1067

```

Db 61 LRTCAHYQNGTLGPDWMAARRKQOLK-----CTNVPRCKTKNAV 103
Qy 500 LLLLTANQTLMGOTSDQVMDKDFLYCAN--APLTFKDGDTALASCMDYAGPVPFLAIG 557
Db 104 YQIL-----HYLVKDKDFWPKT-----ADYATPALKYSMFLP 134
Qy 558 GYKGDYSEALWTSIANNYPAGDRLQAQKLWEAFLEEMRAFORWAGHFOVTTFA 617
Db 135 SPTEKGS-----MMNYLDNF-----ENWNS-----DGVTTIT-- 164
Qy 618 ERSLEDEINTTAEDLPFATSYIVIFLYISLALGSYSWSR--VMVDSKATLGLGVAV 675
Db 165 --GIEFGIKSLFDQYLLMDTVPAIAIVILVVMCVYTKSMFILTMTFAISSLIVSY 222
Qy 676 VLGAVMAGFFSYGLRSSLVLOVVPFLVLSVGADNIFIF--VLEYORLPRRPGPRE 733
Db 223 FLYRVVPHFFPPMNL-TALIIL-----VIGADDAFVLCVDMVNYTKFDPHAETSE 274
Qy 734 VHIGRALGRVAPSMLLCSLEAICFFLGALTPMPAVRTFALTSGLAVIDFLLOMFAVA 793
Db 275 T-VSITLQHAALSMTFTSFTAAAFYANYVSNITAIRCFGVYAGTALVNYVLMVTLPA 333
Qy 794 LLSLDSKROASRLDVCCV-KQBELPPPGOG-----EGLILG-----FPOKAY 836
Db 334 VVVL-----HERYLLNFTCFKKQOQIYDNKSCWTACQKCHKVLPFAISEARIFFEKV 389
Qy 837 APFLHWTIRGVLLFLALFGVLSYMSCHISVGLDQELALPK-----DS 881
Db 390 PCVIRKF--RYLWLFALTVGGAYIVC-----INPKMLPSLELSEFOVFRSSHPFER 442
Qy 882 YLDY-FLFNRYFEVGAPVYVTTILGYNFSE-----AGNATCSSA 923
Db 443 YDAEYKGLFMRVHGEELHMPITVINGVSPEDNGNPLNPKSGKGLTLDSSFNIPAS 502
Qy 924 GCNFSFTOKIYATER--PEOSYLAIPASSWDDFDLWLTSSC-----CRLYISG 973
Db 503 QAWILHFCQKLRNQTFFYQDEQDF-----TSCFIETFKQWENQDCDEPALYPCCHWSF 558
Qy 974 PNKDFPCPSTVNSLNLKNCMSITGMSVRPSVQFHKLPFWFLNDRPNKCPKGGLAAYS 1033
Db 559 PYKQ-----RIFE-----LCIKRAIMELE 577
Qy 1034 TSVNLTSQVLAASRMAYHKLNSQDYTEALRAA-----REL 1072
Db 578 RSTGYHLDSKTPGRF-----DINDTIRAVVLEFQSTYLETLAYEKMHQFYKEV 626
Qy 1073 AMNITADLRKVP-GTDPAPRPVPTITNV-FYEQYLTLPEGLFMLSCLVPTFAVSCLL 1130
Db 627 DSWISSELSAPGLSNG-----WFSNLEFYDLQDSLSDGTLIANGLSVAVAFSV-MLL 680
Qy 1131 LGLDLSGLNLLSIVMILVDTVGFMAWDISYNAVSLNLVSAVGMSVEFVSHITRSPA 1190
Db 681 TTWNIISLYALISIAGTIPVTGSLVLLGWELNLVLESVTSIVAVGLSVDFAHYGVAVR 740
Qy 1191 ISTKPTWLERAKAATISM--GSAVFAVAMTNLPGILVGLAKAQLIQIFFRLNLLIT 1247
Db 741 LAPDP--DREGKVIFSLRSGSAMAMAALTTPVAGAMWMPSTVLAAYTOLGTFM--MLIM 795
Qy 1248 LLGLLHGLVPLVILSVGP 1267
Db 796 CISWAPATFFQCMCRCLGP 815

RESULT 92

US-10-311-623-3
; Sequence 3, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Danniell G.
; APPLICANT: TANG, Y. Tom; LAU, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda

Query Match 2.9%; Score 202; DB 4; Length 1124;
Best Local Similarity 18.9%; Pred. No. 1.5e-07;
Matches 161; Conservative 134; Mismatches 331; Indels 228; Gaps 35;
Qy 519 KDFLYCANAP-----LTPKDGDTALASCMDYAGPVPFLAIGVYKGD 563
Db 6 KQ-LACTNVPRCKTKYNAVYQILHYLVKDKD---FMTPKTADYATPALKYSMFLSPTEKG 61
Qy 564 YSEAEALINTFSILNYPAGDRLQAQKLWEAFLEEMRAFORWAGHFOVTTFAERSLED 623
Db 62 ES-----MMNYLDNF-----ENWNS-----DGVTTIT--GIEF 89
Qy 624 EINRTTAEDLPFATSYIVIFLYISLALGSYSWSR--VMVDSKATLGLGVAVLGVAVM 681
Db 90 GIKHSLFDQYLLMDTVPAIAIVILVVMCVYTKSMFILTMTFAISSLIVSYFLYRVV 149
Qy 682 AAMGPSYLGIRSSLVLOVVPFLVLSVGADNIFIF--VLEYORLPRRPGPREVHIGRA 739
Db 150 FHFEPPFMNL-TALIIL-----VIGADDAFVLCVDMVNYTKFDPHAETSET-VSIT 200
Qy 740 LGRVAPSMLLCSLEAICFFLGALTPMPAVRTFALTSGLAVIDFLLOMSAFVALLSLDS 799
Db 201 LQHAALSMTFTSFTAAAFYANYVSNITAIRCFGVYAGTALVNYVLMVTLPAVVVL-- 258
Qy 800 KRQEASRLDVCCV-KQBELPPPGOG-----EGLILG-----PFQAYAPFLH 842
Db 259 --HERVLLNFTCFKKQOQIYDNKSCWTACQKCHKVLPFAISEARIFFEKVLPVIX 316
Qy 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPK-----DSYLLDY-- 886
Db 317 P--RYLWLFALTVGGAYIVC-----INPKMLPSLELSEFOVFRSSHPPERYDAEYK 369
Qy 887 FLFNRYFEVGAPVYVTTILGYNFSE-----AGNATCSSAGCNFNS 929
Db 370 KLFMFERVHGEELHMPITVINGVSPEDNGNPLNPKSGKGLTLDSSFNIPASQAWILH 429
Qy 930 FTQKIYATER--PEOSYLAIPASSWDDFDLWLTSSC-----CRLYISGPNKDF 979
Db 430 FCQKLRNQTFFYQDEQDF-----TSCFIETFKQWENQDCDEPALYPCCHWSFPYKQ-- 483
Qy 980 CPSTVNSLNLKNCMSITGMSVRPSVQFHKLPFWFLNDRPNKCPKGGLAAYSTSVNLT 1039

Db	484	-----EIFE-----LCIKRAIMELERSTGYH 504	
Qy	1040	SDGOVLASRFMAYHKPLKNSODYTEALRAA-----RELAANITA 1078	
Db	505	LDSTPGPRF-----DINDTIRAVLEFQSTYLFILAYEKHOFYKEVDSWISS 553	
Qy	1079	DLRKVP-GTDPAPAFVFPYITNV-FYEQVLTILPEGLFMLSCLVPTFAVSCLLIGLDLR 1136	
Db	554	ELSSAPEGLSNG-----NFVSNLEFYDLQDSLSDGTLLIAMGLSVAVAFSV-MLLTTWNII 607	
Qy	1137	SGLMLLSIVMLVDTVGFMAWDISYNAVSLINLVSAVGSVERFVSHITSFALSTKPT 1196	
Db	608	ISLYAIISITAGTIFVTGSLVLLGWEANLVESVTISVAVGLSVDPAVHYGAYRLADDP- 666	
Qy	1197	WLERAKEATISM--GSAVFAGVAMTNLPGLILVLGLAKAQLIQIIPFFRLNLIITLLGLH 1253	
Db	667	--DREGKIFLSRVSASAMAAITTFVNGAMMFSTVLAYTQLGTFM--MLINCISWAF 722	
Qy	1254	GLVFLFVILSVYGP 1267	
Db	723	ATFFQCMCRCLGP 736	
RESULT 93			
US-10-085-198-168			
; Sequence 168, Application US/10085198			
; Publication No. US20040009907A1			
; GENERAL INFORMATION:			
; APPLICANT: Alsobrook et al.			
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-279			
; CURRENT APPLICATION NUMBER: US/10/085,198			
; CURRENT FILING DATE: 2002-02-25			
; PRIOR APPLICATION NUMBER: 60/271,646			
; PRIOR FILING DATE: 2001-02-26			
; PRIOR APPLICATION NUMBER: 60/276,401			
; PRIOR FILING DATE: 2001-03-16			
; PRIOR APPLICATION NUMBER: 60/311,981			
; PRIOR FILING DATE: 2001-08-13			
; PRIOR APPLICATION NUMBER: 60/312,858			
; PRIOR FILING DATE: 2001-08-16			
; PRIOR APPLICATION NUMBER: 60/271,840			
; PRIOR FILING DATE: 2001-02-27			
; PRIOR APPLICATION NUMBER: 60/277,324			
; PRIOR FILING DATE: 2001-03-20			
; PRIOR APPLICATION NUMBER: 60/286,096			
; PRIOR FILING DATE: 2001-04-21			
; PRIOR APPLICATION NUMBER: 60/299,695			
; PRIOR FILING DATE: 2001-06-20			
; PRIOR APPLICATION NUMBER: 60/315,614			
; PRIOR FILING DATE: 2001-08-29			
; PRIOR APPLICATION NUMBER: 60/272,405			
; PRIOR FILING DATE: 2001-02-28			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 653			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 168			
; LENGTH: 1561			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-085-198-168			
Query Match 2.9%; Score 197.5; DB 4; Length 1561;			
Best Local Similarity 19.5%; Pred. No. 6e-07;			
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;			
Qy	576	LNNYPAGDPRIAQAUKLWEAFLEEMRAFORMAGMFOVTTAERSLED----- 623	
Db	574	LPNYQSVDDR-----WEE-----ORAKFQSFVVTYVAMLAKOSTSKVQVLYGGTD 618	
Qy	624	----EINRTAED--LPFATSYIVIFLYTISLALGSSYSSRVMVSDSKATLGLGG-VAVV 676	
Db	619	LFDYEVRTFNMDMLAFISSCIAALVILTSCSVLSFPGI-----ASIGLSCLVALF 673	

RESULT 94
US-10-239-316-17
; Sequence 17, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun

Qy	677	LGAVMAANGPFSYLGIRSSLVILQVVPFLVLSVCADNIFIVLVEYQRLPR-RPCEPREVH 735	
Db	674	LYHVVFGI---QVYLGLNG-----VAAPFVIGVDDVDFVINTYRQATHLEDQLRMH 725	
Qy	736	IGRALGRVAPSMILCSLSEACFPLGALTMPAVRTPALTSGLAVI---LDLFQMSAFV 792	
Db	726	TVQTAGK---ATFPTSITTAAYAAVPSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAL 782	
Qy	793	ALLSLDSKROEASRLDVC--CCVKPQBLPPPGQ----- 823	
Db	783	GLWSLYLAPLESSQTSCHQNSRKTSLHPFGDFVATPEQVGGSPAQAPIFYLDDDIPLL 842	
Qy	824	-----GEGLLGLFFQKAYAP-----FLLHWT-----TRGVVL 850	
Db	843	EVBEPEVSLGLDVSLSVSPEGLQPASNTSGRGLHVLQLOELLHHWVAWKSRRWIV 902	
Qy	851	LLFLALFGVSL-----YSM-----CHISVGLDQE- 874	
Db	903	GLFVSIILSLVPASRLRPASRAPLLPRPDNIQVLLDLKYNLSAEGISCITCSLFOBK 962	
Qy	875	-----LALPK-----DSYLLDY--FLFLNRYFEVGAPVFTVLGYNF 910	
Db	963	PHSLQNNIRTSLEKKRSGSGVPMASRPEATLQDFPGTVIYSKVKSQGHPA--VYRLSLNA 1020	
Qy	911	SSEAGMNAICSSAG-CNNF-----SFTOKIQVATEFPEQSVYLAIAPASSWVDDFIDLW 961	
Db	1021	SLPAPWQAVSPGDGEVPSFQVYRAPFGNFTKKLTACMSTVGLLQNASPSRKM- -L 1074	
Qy	962	TPSSC-----CRLYIS--GP 974	
Db	1075	TTLACDAKRWKDFDFYVATKEQHTKLYFAQSHKPPFHGRVCMAPPGCLLSSPDGP 1134	
Qy	975	NKDFCPSPTVNSLNCNKNCMTMG-----SVRPSVEGFHKYLPWF----LN- 1017	
Db	1135	TKGFF---FVPSEKVPKARLSATFGFNPFCVNTGCGKPAVRPLVDTGAMVVFVFGIIGNR 1191	
Qy	1018	-----DRPNIKCPKGLAAYSTSVNLTSD-----GOVL 1045	
Db	1192	TRQVDNHHVIGDP--GSVYVDSFDLPKEIGHLCHLCKAIAANSELVKPGCAQLPSGYSI 1249	
Qy	1046	ASRFMAYHKPLKNSQD-----YTEALRAARELAANITADLRKVPGTDPAPFV-P 1094	
Db	1250	SSFQWLHPECKELPEPNLLPQLSHGAVGVRGREGVQVEISMASESTTYKGS-SFQTSYSD 1308	
Qy	1095	YTITNVFYEQYLTILPEG-----LPM-----LSLCL--VP 1122	
Db	1309	YLRWESFLQQQLQALPEGSVLRRGFTCEHWKQIFMEIVGVQSALCGLVLSLLICVAVA 1368	
Qy	1123	TFAVSCLLGLDLRSGLNLSIVMLVDTVGFMAWDISYNAVSLINLVSAVGSVEFV 1182	
Db	1369	VFTTHILL-----LLPVLVLSILGIVCLLV--TIMYWSGEM--GAVEAISILVGSVDYC 1421	
Qy	1183	SHITRSFALS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI----- 1224	
Db	1422	VHLVEGYLLAGENLPHPQAEADARTQRM--RTLEAVRHVGVAIVSSALTTVATVPLFFC 1479	
Qy	1225	LVLGLAKAQLIQIFPPRLMLLITLLGLHGLFPLPVILSVGVDPVNPALAEQKRAE--E 1282	
Db	1480	IIAPPAK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTKRTISFLK 1527	
Qy	1283	AVAAVMVA 1290	
Db	1528	ALGAVLLA 1535	

```
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same An
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-17

Query Match      2.8%; Score 195.5; DB 4; Length 1087;
Best Local Similarity 19.5%; Pred. No. 5.1e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTTAERSLED-----623
Db 100 LPNYYSVDDR-----WEE-----QRAKFSFVTVVAMLAQSTSKVQVLYGGTD 144
QY 624 -----EINRTAED--LPFATSYIVIFLYISALGYSYSSWSRVWVDSKATLGLGG-VAVV 676
Db 145 LFDYEVRRFTFNNDMLAFITSSCIAALVYILTSCSVFLSFFGI-----ASIGLSCLVALF 199
QY 677 LGAVMAAMGFFSYLGRSSVLIVQVVPFLVLSVGADNIFIVLEXYQRLPR-RPGEPREYH 735
Db 200 LTHVFGI---QYLGILNG-----VAAFVIGVIGVDDVVFVINTYRQATHLEDPPQLRMTH 251
QY 736 IGRALGRVAPSMLLCSLSEAICFFFLGALTMPAVRTFALTSLGLAVI---LDFLLOMSAFV 792
Db 252 TWQTAGK---ATFTSLTTAAAYAAANVFQIIPAVHDFGLFMSLVSCCVLAVLVTMPAAL 308
QY 793 ALLSDSKRQEARLDVC--CCVKQDELPPGQ-----823
Db 309 GLWSLYLAPLESSCQTSCHQNSRKTSLHFPDGVFAAPQVGGSPAQGPYPYLDLDDIPL 368
QY 824 -----GEGLLGFFQKAYAP-----FLLHWI-----TRGVVL 850
Db 369 EYVEEPVSLLELGDVLSVSPGELQPAASNTGSRGHLIVQLQELHHWLSAVKSRWIV 428
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
Db 429 GLFVSLIILSLVPASRLRPASRAPLLFRPDTHIQVLLDLKYNLSAEGISCITCSGLFQBK 488
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFEVGPVYFVTTLYGYNF 910
Db 489 PHSLQNNIRTSLEKKRRSGVPMASRPEATLQDPFGTVVISKVKSQHFA--VYRLSLNA 546
QY 911 SSEAGNNAICSSAG-CNPF-----STQKIQYATEPPEOSYLAI PASSWDDFDWL 961
Db 547 SLPAWQAVSPGDGEVPFQVYRAPFGNFTKXLTACMSTVGLLQAASPRKWM-----L 600
QY 962 TPSSC-----CRLYTS--GP 974
Db 601 TTLACDAKRGKWFDFSYVATKEQHQTRKLYPAQSHKPPFHGRVCMAPPGCLLSSSPDG 660
QY 975 NKDKFCFSTVNSLNLKNCWSITMG-----SVRPSVEQPHKYLPMF-----LN- 1017
Db 661 TKGFF---FVPSKVPKARLSATFGFNPVNTGCGKPAVRPLVDTGAMVYFVFGIIGVNR 717
QY 1018 ---DRPNIKCPKGGLAAYSTSVNLTSD-----GQVL 1045
Db 718 TRQVDNHVIGDP--GSVYDSSFDLFKEIGHLCHLCKAIAANSSELVKPGGAQCLPSGYSI 775
QY 1046 ASREPMAYHKPLKNSQD-----YTEALRAARELANITADLRKVPDTPAPEVP-P 1094
Db 776 SSFLQMLHPECKELPEPNLLPQOLSHGAVGVRGVQWISMAFESTYTKGKS-SFQTSYSD 834
QY 1095 YTITNVFVEQYLTILPEG-----LFW-----LSLCL--VP 1122
```

```
Db 835 YLRWESFLQOQLOALPEGSVLRGRFQTCBHWKQIFMEIVGVQSALCGLVSLICVAANA 894
QY 1123 TPAVSCLLGLDLRSGLNLLSTVMILVDTVGFMALWDISYNAVSLINLVSAGVMSVEFV 1182
Db 895 VFTTHILL-----LLPVLSILGIVCLV--TIMYWSGEM--GAVEAISLSILVSSVDYC 947
QY 1183 SHITRSPAIS-----TKPTWLERAKBATISMGSAVFAGVAMTNLPGI-----1224
Db 948 VHLVEGYLLAGENLPHQAEADARTQW--RTLEAVRHVGVAIVSSALTTVIAVPLPFFC 1005
QY 1225 LVVLGAKAQLIQIFPFRNLNLLITLLGLHGLVPLVILSVGVDPVNPALALEQKRAE--E 1282
Db 1006 IIAFPK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTRTTSFLK 1053
QY 1283 AVAAVMVA 1290
Db 1054 ALGAVLLA 1061

RESULT 95
US-10-239-316-40
; Sequence 40, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same An
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 40
; LENGTH: 1392
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-40

Query Match      2.8%; Score 195.5; DB 4; Length 1392;
Best Local Similarity 19.5%; Pred. No. 7.5e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTTAERSLED-----623
Db 405 LPNYYSVDDR-----WEE-----QRAKFSFVTVVAMLAQSTSKVQVLYGGTD 449
QY 624 -----EINRTAED--LPFATSYIVIFLYISALGYSYSSWSRVWVDSKATLGLGG-VAVV 676
Db 450 LFDYEVRRFTFNNDMLAFITSSCIAALVYILTSCSVFLSFFGI-----ASIGLSCLVALF 504
QY 677 LGAVMAAMGFFSYLGRSSVLIVQVVPFLVLSVGADNIFIVLEXYQRLPR-RPGEPREYH 735
Db 505 LTHVFGI---QYLGILNG-----VAAFVIGVIGVDDVVFVINTYRQATHLEDPPQLRMTH 556
QY 736 IGRALGRVAPSMLLCSLSEAICFFFLGALTMPAVRTFALTSLGLAVI---LDFLLOMSAFV 792
Db 557 TWQTAGK---ATFTSLTTAAAYAAANVFQIIPAVHDFGLFMSLVSCCVLAVLVTMPAAL 613
QY 793 ALLSDSKRQEARLDVC--CCVKQDELPPGQ-----823
Db 614 GLWSLYLAPLESSCQTSCHQNSRKTSLHFPDGVFAAPQVGGSPAQGPYPYLDLDDIPL 673
QY 824 -----GEGLLGFFQKAYAP-----FLLHWI-----TRGVVL 850
Db 674 EYVEEPVSLLELGDVLSVSPGELQPAASNTGSRGHLIVQLQELHHWLSAVKSRWIV 733
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
```

Db 734 GLFVILSLVFAASRLRASPAPLLFRPDNNIQVLLDLKYNLSAEGISCITCSGLFOEK 793
Qy 875 -----LALPK-----DSYLLDY--FLFLNRYFEVGPVVFVTLGLGYNF 910
Db 794 PHSLQNNRTSLKERRGSGVPWASRPENTLQDFPGTVYISKVKSQGHPA--VYRLSLNA 851
Qy 911 SSEAGNAICSGAG--CNPF-----SFTQKIQVATEFPEQSYLAIPASSWDDFDIWL 961
Db 852 SLPAQWAVSPGDGVPSPFQVYRAPFGNFTKLTACMSTVGLLQAASPSRKM-----L 905
Qy 962 TPSSC-----CRLYTS--GP 974
Db 906 TTLACDAKRGWKFDPSFYVATKEOQTRKLYPAQSHKPPFHGRVCMAPPGCLSSSPGP 965
Qy 975 NKDFCPSVNSLNCNKQMSITMG-----SVRPSVEQFHLYLPMF-----LN- 1017
Db 966 TKGFF--FVPESEKPKARLSATFCGNPCVNTGCGKPAVRPLVDGMVFFVFGIIGVNR 1022
Qy 1018 ----DRPNIKCPKGLAAYSTSVNLTS-----QVQL 1045
Db 1023 TRQVDNHVIGDP--GSVVYDSSFDLFKEIGHLCHLCKAIAANSSELVKPGGAQLPSGYSI 1080
Qy 1046 ASRFWAYHKPLKNSQD-----YTEALRAARELAANITADLRKVPGTDPAPFVP-P 1094
Db 1081 SSFLOMLHPECKELPEPILLPQOLSHGAVGREGRVQWISMAFESTTYKGS--SFQTSYSD 1139
Qy 1095 YTITNVFYQYLTIPEG-----LFM-----LSLCL--VP 1122
Db 1140 YLRWESFLOQQLALPEGSVLRGRGTCEHMKQIFMEIVGVQSALCGLVLSLICVAVA 1199
Qy 1123 TFVASCLLGLDLRGLNLLSIIVMLVDTVGFMALNDSYNVNSLINLVSAGHVSVEFV 1182
Db 1200 VETTHILL----LLPVLISILGIVCLAV--TIMYWSGWM--GAVEAISLSILVGSVDYC 1252
Qy 1183 SHITRSFAPS-----TKPTMLERAKEATISMGSAVFAGVAMTNLPGI----- 1224
Db 1253 VHLVEGYLLAGENLPPHQADARTQW--RTLEAVRHVGVAIVSSALTVIATVPLFPFC 1310
Qy 1225 LVGLAKAQIOLIFFPRLNLLITLLGLHLGLVLPVILSVYVGDVNPALALBQKRAE--B 1282
Db 1311 ITAPPAK-----FGKIVALNTGVSIYTLTVSTALLGIMAPS-----SFTTRTTSFLK 1358
Qy 1283 AVAANMVA 1290
Db 1359 ALGAVLLA 1366
RESULT 96
US-10-437-963-102679
; Sequence 102679, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 102679
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100179C.1.pgp
US-10-437-963-102679

Query Match 2.8%; Score 190; DB 4; Length 100;
Best Local Similarity 41.2%; Pred. No. 3.7e-08;
Matches 33; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
Qy 1200 RAKEATISMGSAVAGVAMTNLPGILVLGLAKAQIOLIFFPRLNLLITLLGLHLGLVLP 1259
Db 15 RARQALSTMGASVSGITLTKLVGVIVLRPAKSEVFFVYVQMYLALVIGFLHGLIFLP 74
Qy 1260 VILSVYVGDVNPALALEQKR 1279
Db 75 VVLSLCGPPSKVMKPLEQSQ 94
RESULT 97
US-10-168-428-3
; Sequence 3, Application US/10168428
; Publication No. US20030165897A1
; GENERAL INFORMATION:
; APPLICANT: Universit,t Zrich
; TITLE OF INVENTION: Dispatched Polypeptides
; FILE REFERENCE: D. melanogaster dispatched sequence
; CURRENT APPLICATION NUMBER: US/10/168,428
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 09/468,237
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-168-428-3
Query Match 2.7%; Score 187.5; DB 4; Length 931;
Best Local Similarity 20.2%; Pred. No. 2e-06;
Matches 152; Conservative 119; Mismatches 252; Indels 231; Gaps 36;
Qy 606 RMAG-----MPQVTTFAERSLEDEINRTTAEDLPATSYI-----VIFLYISLALGS 653
Db 299 RMAGIALANMKNKVFQ-----ERIQTDSLFAAFSA--LLVPSCFLIYSRSIIFTCIILMVVT 352
Qy 654 YSSWSRVWVDSKATLGLGVAVVLGAVMAAMGFFSVGLIRSSLVILQVVPFLVLSVGADN 713
Db 353 LSS-----GVAFITVVLGIDFFPINL-----LVVILISIGADD 389
Qy 714 IFIFVLEYQRLPRRPFCEPREVHIGR-----ALGRVAPSMMLLSLSEA 755
Db 390 AFLLLVYRREVERMSH--LEYKVGSIYIPLYRESOLLSRSLRSLHSHLSVSMFVTSLTTA 448
Qy 756 ICFFLGLATPMPAVRTFALTSLGLAVILDLLOWSAFVALLSLDSK--RQASRLDVCCKVK 814
Db 449 STFLTNLSSPVLRLCFGYAALTVTNVILVVLIPGAILLSRPIRKKLSRGD-----EE 504
Qy 815 POELPPPGQEGLLGLFFOKAVAPF-----LLHWITRGVLLLLFLALFGVSLYSMCHIS 868
Db 505 PEKI-----ESHYFASKITETHTHYRFG--IFICSLIMTGLSLFIIFQ-N 547
Qy 869 VGLDQELALPK-----DSYLLDYF-----LFLNRYFEVGPVVFVTLGLY 908
Db 548 PGLKTPQTNPTKLLVDSNTHIEYFDNNVHFNFQWQSARLVKNFVFGVDAIKETSLSPY 607
Qy 909 NFSSEAGMAICSSACCNFS-----FTQKIQY---ATEFPQSQSLAIPAS--SWWDDF 957
Db 608 NKFSK-----NFSQAHYSLDTDKLDYFRRIVNLSKSKYQLVNYTHVSWADKI 654
Qy 958 IDWLTPSSC-----CRLYISGNPKDKFCPSTVNSLNCNKQMSITMGSRVPSVEQ 1007
Db 655 LQ--ANESCFSENKTIHCEILSASVRNK-----NLHQFPDDFSVIPGD--GPIDQ 703
Qy 1008 FHKYLEPWLNDNRNIKCPKGGGLAAYSTSVNLTSVGQVLARFMAYH---KPLKNSQDYTE 1064
Db 704 DLKVVGYFISIPSNQK-----LQVDTEMIGSFQFQIEBSCKQIKNATSDSV 749

Tue Apr 11 10:57:57 2006

us-10-736-769-4.top100.rapbm

```
QY 1187 RSPAISTKP---TWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFFRLN 1243
Db 966 GNYCTDKKPAINWIE-----GRGKSV---VTEAVIPGAIVEKVLKTTVAAL--VELN 1013
QY 1244 LLITLLGL-----LHGLVFLPVILSVGPDVNPALALEOKRAEEAAVAV 1287
Db 1014 ISKNLIGSAMAGSVGGFNAHAANILTAIYLATGQD--PAQNVSSNCITLMKAV 1065
```

```
RESULT 100
US-10-060-756A-14
; Sequence 14, Application US/10060756A
; Publication No. US20030046717A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/060,756A
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4804
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 14
; TYPE: PRT
; LENGTH: 101
; ORGANISM: Homo sapiens
US-10-060-756A-14
```

```
Query Match 2.5%; Score 170; DB 4; Length 101;
Best Local Similarity 36.0%; Pred. No. 2.1e-06;
Matches 36; Conservative 26; Mismatches 36; Indels 2; Gaps 2;

QY 1167 SLINLVSAGCMSEVFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILV 1226
Db 1 SMINLVICTGFDFDFAHISYAFVSSSQPSVQKSVKVEALYLLGYPVLQS-AISTIGVCV 59

QY 1227 LGLAKAQLIQIFFFRLNLLITLLGLHGLVFLPVILSVYG 1266
Db 60 LAAAKAYIPT-FFKIMFLVMIFGAHGLIFIPVLTFFG 98
```

Search completed: April 11, 2006, 01:01:15
Job time : 232 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:58:02 ; Search time 27 Seconds
(without alignments)
1538.804 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLRLAQ.....GSIKGAGATNPLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications_AA_New:*
1: /SIDSS/ptodata/2/pubppaa/US06_NEW_PUB.pdb:*
2: /SIDSS/ptodata/2/pubppaa/US06_NEW_PUB.pdb:*
3: /SIDSS/ptodata/2/pubppaa/US07_NEW_PUB.pdb:*
4: /SIDSS/ptodata/2/pubppaa/PCT_NEW_PUB.pdb:*
5: /SIDSS/ptodata/2/pubppaa/US09_NEW_PUB.pdb:*
6: /SIDSS/ptodata/2/pubppaa/US10_NEW_PUB.pdb:*
7: /SIDSS/ptodata/2/pubppaa/US11_NEW_PUB.pdb:*
8: /SIDSS/ptodata/2/pubppaa/US60_NEW_PUB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6896	99.8	1332	7	US-11-242-459-9
2	2402.5	34.8	1278	6	US-10-995-561-952
3	942	13.6	194	7	US-11-242-459-8
4	195.5	2.8	1087	7	US-11-242-459-17
5	195.5	2.8	1392	7	US-11-242-459-40
6	191	2.8	127	7	US-11-096-568A-6778
7	173	2.5	85	7	US-11-096-568A-6779
8	143	2.1	445	7	US-11-242-459-35
9	143	2.1	456	7	US-11-242-459-34
10	141.5	2.0	1048	6	US-10-392-234A-18
11	140.5	2.0	1046	6	US-10-392-234A-16
12	134	1.9	1067	6	US-10-467-657-2648
13	127	1.8	1048	6	US-10-392-234A-14
14	126.5	1.8	200	7	US-11-242-459-16
15	126	1.8	801	6	US-10-793-626-2020
16	121.5	1.8	1043	6	US-10-392-234A-34
17	120.5	1.7	2426	7	US-11-203-806A-11
18	119	1.7	1048	6	US-10-392-234A-20
19	117.5	1.7	1049	6	US-10-392-234A-12
20	115.5	1.7	914	7	US-11-072-512-2923
21	113	1.6	1034	6	US-10-392-234A-30
22	112.5	1.6	597	7	US-10-392-389-362
23	109.5	1.6	758	7	US-11-096-568A-30412
24	109.5	1.6	772	7	US-11-096-568A-30411
25	109.5	1.6	827	7	US-11-096-568A-30410

1049	1.6	109.5	26	7	US-11-098-686-11239	Sequence 11239, A
1032	1.6	109	27	6	US-10-392-234A-67	Sequence 67, Appl
648	1.6	108.5	28	7	US-11-087-099-4135	Sequence 4135, Ap
443	1.5	106	29	7	US-11-087-099-11457	Sequence 11457, A
783	1.5	105.5	30	7	US-11-082-389-354	Sequence 354, App
324	1.5	105	31	7	US-11-087-099-11792	Sequence 11792, A
1080	1.5	105	32	7	US-11-096-568A-27723	Sequence 27723, A
1097	1.5	105	33	7	US-11-096-568A-27722	Sequence 27722, A
1128	1.5	105	34	7	US-11-096-568A-27721	Sequence 27721, A
315	1.5	104.5	35	7	US-11-087-099-9820	Sequence 9820, Ap
471	1.5	104.5	36	7	US-11-087-099-4362	Sequence 4362, Ap
757	1.5	104	37	7	US-11-096-568A-27509	Sequence 27509, A
771	1.5	104	38	7	US-11-096-568A-27508	Sequence 27508, A
826	1.5	104	39	7	US-11-096-568A-27507	Sequence 27507, A
251	1.5	103.5	40	7	US-11-087-099-1098	Sequence 1098, Ap
331	1.5	103.5	41	7	US-11-087-099-8831	Sequence 8831, Ap
2042	1.5	103.5	42	6	US-10-506-454-134	Sequence 134, App
2312	1.5	103.5	43	7	US-11-126-313-34	Sequence 34, Appl
463	1.5	103	44	7	US-11-087-099-3980	Sequence 3980, Ap
475	1.5	102.5	45	7	US-11-087-099-2723	Sequence 2723, Ap
324	1.5	102	46	7	US-11-087-099-5717	Sequence 5717, Ap
647	1.5	101	47	7	US-11-087-099-9244	Sequence 9244, Ap
726	1.5	100.5	48	7	US-11-096-568A-2446	Sequence 2446, Ap
727	1.5	100.5	49	7	US-11-096-568A-2445	Sequence 2445, Ap
836	1.5	100.5	50	7	US-11-096-568A-2444	Sequence 2444, Ap
279	1.4	100	51	7	US-11-087-099-4795	Sequence 4795, Ap
3389	1.4	100	52	6	US-10-204-252-10	Sequence 10, Appl
3391	1.4	100	53	6	US-10-204-252-6	Sequence 6, Appl
3391	1.4	100	54	6	US-10-204-252-8	Sequence 8, Appl
3391	1.4	100	55	6	US-10-204-252-12	Sequence 12, Appl
3391	1.4	100	56	6	US-10-204-252-16	Sequence 16, Appl
3391	1.4	100	57	6	US-10-204-252-28	Sequence 28, Appl
577	1.4	99.5	58	6	US-10-718-264-3	Sequence 3, Appl
577	1.4	99.5	59	6	US-10-718-264-3	Sequence 3, Appl
716	1.4	99	60	6	US-10-506-454-1056	Sequence 1056, Ap
3391	1.4	98	61	6	US-10-204-252-14	Sequence 14, Appl
3402	1.4	98	62	6	US-10-204-252-18	Sequence 18, Appl
324	1.4	97	63	7	US-11-087-099-4100	Sequence 4100, Ap
336	1.4	96.5	64	7	US-11-087-099-7384	Sequence 7384, Ap
465	1.4	96.5	65	6	US-10-793-626-2928	Sequence 2928, Ap
836	1.4	96	66	6	US-11-087-099-2386	Sequence 2386, Ap
294	1.4	96	67	6	US-10-858-730-118	Sequence 118, App
479	1.4	96	68	7	US-11-087-099-4490	Sequence 4490, Ap
479	1.4	96	69	7	US-11-087-099-10324	Sequence 10324, A
479	1.4	96	70	7	US-11-087-099-11713	Sequence 11713, A
983	1.4	96	71	6	US-10-511-538-91	Sequence 91, Appl
1167	1.4	96	72	6	US-11-097-125-2	Sequence 2, Appl
338	1.4	95.5	73	7	US-11-087-099-8709	Sequence 8709, Ap
686	1.4	95.5	74	6	US-10-821-234-1027	Sequence 1027, Ap
2186	1.4	95.5	75	7	US-11-096-568A-28283	Sequence 28283, A
2199	1.4	95.5	76	7	US-11-096-568A-28282	Sequence 28282, A
2301	1.4	95.5	77	7	US-11-096-568A-28281	Sequence 28281, A
325	1.4	95	78	7	US-11-087-099-9929	Sequence 9929, Ap
461	1.4	95	79	7	US-11-183-205-32	Sequence 32, Appl
461	1.4	95	80	7	US-11-260-192-2	Sequence 2, Appl
469	1.4	95	81	7	US-11-087-099-11321	Sequence 11321, A
552	1.4	95	82	7	US-11-087-099-10983	Sequence 10983, A
758	1.4	95	83	6	US-10-505-263-8	Sequence 8, Appl
793	1.4	95	84	7	US-11-087-099-2968	Sequence 2968, Ap
231	1.4	95	85	7	US-11-087-099-4589	Sequence 4589, Ap
467	1.4	94.5	86	7	US-11-087-099-2054	Sequence 2054, Ap
467	1.4	94.5	87	6	US-10-479-873-9	Sequence 9, Appl
599	1.4	94.5	88	7	US-11-087-099-5936	Sequence 5936, Ap
336	1.4	94	89	7	US-11-205-109-4	Sequence 4, Appl
470	1.4	94	90	6	US-10-511-538-83	Sequence 83, Appl
1431	1.4	94	91	7	US-11-128-059-2	Sequence 2, Appl
296	1.4	93.5	92	7	US-11-082-389-16	Sequence 16, Appl
461	1.4	93.5	93	7	US-11-087-099-2325	Sequence 2325, Ap
461	1.4	93.5	94	7	US-11-087-099-4376	Sequence 4376, Ap
461	1.4	93.5	95	7	US-11-087-099-9669	Sequence 9669, Ap
602	1.4	93.5	96	7	US-11-087-099-10340	Sequence 10340, A
679	1.4	93.5	97	6	US-10-506-454-279	Sequence 279, App

99	93.5	1.4	824	7	US-11-090-617-657	Sequence 657, App
100	93.5	1.4	4868	7	US-11-044-111-24	Sequence 24, App1
ALIGNMENTS						
RESULT 1						
US-11-242-459-9						
; Sequence 9, Application US/11242459						
; Publication No. US20060035835A1						
; GENERAL INFORMATION:						
; APPLICANT: TANIYAMA, Yoshio						
; APPLICANT: KITA, Shunbun						
; APPLICANT: SATOMI, Tomoko Komiyama						
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof						
; FILE REFERENCE: 2703USOP						
; CURRENT APPLICATION NUMBER: US/11/242,459						
; CURRENT FILING DATE: 2005-10-03						
; PRIOR APPLICATION NUMBER: US/10/239,316						
; PRIOR FILING DATE: 2002-09-19						
; PRIOR APPLICATION NUMBER: PCT/JP01/02279						
; PRIOR FILING DATE: 2001-03-22						
; PRIOR APPLICATION NUMBER: JP2000-088595						
; PRIOR FILING DATE: 2000-03-24						
; NUMBER OF SEQ ID NOS: 59						
; SEQ ID NO 9						
; LENGTH: 1332						
; TYPE: PRT						
; ORGANISM: Human						
US-11-242-459-9						
Query Match						
Best Local Similarity 99.8%; Score 6896; DB 7; Length 1332;						
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;						
QY	1	MAEAGLRGWL	ALLRLAQSEPYTHI	HOPGYCAFYDECGKPELSGSLMTLSNV	SCLSN	60
DB	1	MAEAGLRGWL	ALLRLAQSEPYTHI	HOPGYCAFYDECGKPELSGSLMTLSNV	SCLSN	60
QY	61	TPARKITG	HLILLOKICRLYTG	PNTOACCSAKQLVLEASISITKALLTR	CPACSDNF	120
DB	61	TPARKITG	HLILLOKICRLYTG	PNTOACCSAKQLVLEASISITKALLTR	CPACSDNF	120
QY	121	VNLHCHNT	CSPNQSLFINTRVA	QLGAGQLPAVVAEAFYQHSFABQSYDSCSRVRPAA	180	
DB	121	VNLHCHNT	CSPNQSLFINTRVA	QLGAGQLPAVVAEAFYQHSFABQSYDSCSRVRPAA	180	
QY	181	ATLAVGTM	CGVYGSALCNAQR	WLNFGDGTGNGLAPLDITPHLLERGOAVGSIQPLNEGV	240	
DB	181	ATLAVGTM	CGVYGSALCNAQR	WLNFGDGTGNGLAPLDITPHLLERGOAVGSIQPLNEGV	240	
QY	241	ARCNESQ	GDVATCSQDCQA	SPAIARPOALDSTFVLGMPGSLVLIILICSVFAVVTI	300	
DB	241	ARCNESQ	GDVATCSQDCQA	SPAIARPOALDSTFVLGMPGSLVLIILICSVFAVVTI	300	
QY	301	LLVGFVR	VPARDKSKMVD	PKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV	360	
DB	301	LLVGFVR	VPARDKSKMVD	PKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV	360	
QY	361	IPVVALA	GLVFTLTDTP	VELMSAPNSQARSEKAFHDQHGFPFRTNQVILTAPNRSY	420	
DB	361	IPVVALA	GLVFTLTDTP	VELMSAPNSQARSEKAFHDQHGFPFRTNQVILTAPNRSY	420	
QY	421	RYDSLIG	PKNFSGILD	DLLELELQERLRLHQLVMSPEAQRNISLQDICYAPLNPDNT	480	
DB	421	RYDSLIG	PKNFSGILD	DLLELELQERLRLHQLVMSPEAQRNISLQDICYAPLNPDNT	480	
QY	481	SLYDCC	INSLLIQYFQNN	RNTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG	TALAL	540
DB	481	SLYDCC	INSLLIQYFQNN	RNTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG	TALAL	540
QY	541	SCMADY	GAPVFPFLA	IGGYKGKDYSEAEALIMTSLNNYPAGDPRLAQAKLWEEAFLEEM	600	

1046 ASRFMAVHKPLKNSQD-----YTEALRAARELAANITADLRKVPOTDPAFEVFP-P 1094
776 SSFLQMLHPECKELPEPNLLPGQLSHGAVGRVQWISMAFESTTYKGS-SFQTYSD 834
1095 YTTITNVFYEQLTILPEG-----LFM-----LSLCL--VP 1122
835 YLWESFLQQLQALPEGSVLRGRFQTCHEHWKQIFMBEIVGVQSALCGLVLSLLICVAAVA 894
1123 TFAVSCLLGLDLRSGLNLSLVMILVDTVGMALWDISYNAVSLINLYSAVGMSVEFV 1182
895 VFTTHILL-----LLPVLLSILGIVCLVV--TIMWSGHEM--GAVEALSLSLTVSSVDYC 947
1183 SHITRSPAIS-----TKPTWLBRAKEATISMGSAPVAGVAMTNLPGI----- 1224
948 VHLVEGYLLAGENLPPHQAEADARTQRM--RTLEAVRHVGVAIVSSALTTVIATVPLPFC 1005
1225 LVLGLAKAQLIQIPFFRLNLLITLLGLHLGLVFLPVLISYVGPDPNPALEQKRAE--E 1282
1006 IIAFFAK-----FGKIVALNTGVSILYTLTVSTALLGINAPS-----SPTRTTSFLK 1053
1283 AVAAVMVA 1290
1054 ALGAVLLA 1061
RESULT 5
US-11-242-459-40
; Sequence 40, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-17
Query Match 2.8%; Score 195.5; DB 7; Length 1087;
Best Local Similarity 19.5%; Pred. No. 6.1e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;
QY 576 LNNYPAGDRLAQAKLWEEAFLEEMRAFORRMAGMFQVTFTAERSLED----- 623
DB 100 LPNYYSVDDR-----WEB-----ORAFQSFVTVYVAMLAQKSTSKVQVLYGGTD 144
QY 624 ----EINRTTAED--LPFATSYVIVFLYISALGSYSSWSRMVDSKATLGLGG-VAVV 676
DB 145 LPDYEVRRFTFNMDLLAFISSCIAALVILTSCSVFLSFFGI-----ASIGLSCLVALP 199
QY 677 LGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLEYOQLRPR-RPGEPREVH 735
DB 200 LYHVVFGEI---QYLGILNG-----VAAFVIVGIGVDDVFVINTYRQATHLEDPLQRMH 251
QY 736 IGRALGRVAPSMLLCSLSEATCFPLGALTTPMPAVRTFALTSGLAVI---LDFLQMSAFV 792
DB 252 TVQTAGK---ATFTSLTTAAVAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308
QY 793 ALLSLDSKQBASRLDVC--CCVKPQELPPPGQ----- 823
DB 309 GLWSLYLAPLESSQTSCHQNCRSKTSLHFFGDVFAAPEQVGGSPAQGPPIYLDLDDIPL 368
QY 824 -----GEGLLGFTQKAYP-----FLHWHI-----TRGVVL 850
DB 369 EVEBEPVSLGLDVSLSVSPGEGIQPASNTGSRGHLIVQLQELHHVWLSAVKSRVIV 428
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
DB 429 GLFVSILILSLVAFASRLRPSRAPLLRPDNTIOVLDLKYNLSAEGISICITCSGLFQEK 488
QY 875 -----LALPK-----DSVLLDY--FLFLNRYPEVGAPVYVFTTLGYNF 910
DB 489 PHSIQNNIRTSLEKRRGSGVPMASRPENTIQDPFGTVYISKVKSQGHPA--VYRLSLNA 546
QY 911 SSEAGMNAICSSAG-CNNF-----SFTQKIQYATEFFEQSYLAIPASSWDDDFDML 961
DB 547 SLPAWQAVSPGDGEVPSFOYRAPFGNFTKCLTACMSTVGLLQAAFPSRKWM-----L 600
QY 962 TPSSC-----CRLYIS--GP 974
DB 601 TTLACDAKRWKPFDSFVATKEOQHTKLYPAQSHKPPFHGRVCMAPPGLLSSSPDGP 660
QY 975 NKDKFCPSTVNSLNCNKMGITMG-----SVRPSVQGFHKYLPWF-----LN- 1017
DB 661 TKGFF---FVPSKYPKARLSATFGFNCVNTGCGKPAVRPLVDGTGMVVFVIGIYNR 717
QY 1018 ----DRPNKCPKPGGLAAYSTVNLTS-----GQVL 1045
DB 718 TRQVDNHWIGDP--GSVVYDSSFDLFKEIGHLCHLCKAIAANSELVKPGGAQCLPSGYSI 775

GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-17
Query Match 2.8%; Score 195.5; DB 7; Length 1087;
Best Local Similarity 19.5%; Pred. No. 6.1e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;
QY 576 LNNYPAGDRLAQAKLWEEAFLEEMRAFORRMAGMFQVTFTAERSLED----- 623
DB 100 LPNYYSVDDR-----WEB-----ORAFQSFVTVYVAMLAQKSTSKVQVLYGGTD 144
QY 624 ----EINRTTAED--LPFATSYVIVFLYISALGSYSSWSRMVDSKATLGLGG-VAVV 676
DB 145 LPDYEVRRFTFNMDLLAFISSCIAALVILTSCSVFLSFFGI-----ASIGLSCLVALP 199
QY 677 LGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLEYOQLRPR-RPGEPREVH 735
DB 200 LYHVVFGEI---QYLGILNG-----VAAFVIVGIGVDDVFVINTYRQATHLEDPLQRMH 251
QY 736 IGRALGRVAPSMLLCSLSEATCFPLGALTTPMPAVRTFALTSGLAVI---LDFLQMSAFV 792
DB 252 TVQTAGK---ATFTSLTTAAVAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308
QY 793 ALLSLDSKQBASRLDVC--CCVKPQELPPPGQ----- 823
DB 309 GLWSLYLAPLESSQTSCHQNCRSKTSLHFFGDVFAAPEQVGGSPAQGPPIYLDLDDIPL 368
QY 824 -----GEGLLGFTQKAYP-----FLHWHI-----TRGVVL 850
DB 369 EVEBEPVSLGLDVSLSVSPGEGIQPASNTGSRGHLIVQLQELHHVWLSAVKSRVIV 428
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
DB 429 GLFVSILILSLVAFASRLRPSRAPLLRPDNTIOVLDLKYNLSAEGISICITCSGLFQEK 488
QY 875 -----LALPK-----DSVLLDY--FLFLNRYPEVGAPVYVFTTLGYNF 910
DB 489 PHSIQNNIRTSLEKRRGSGVPMASRPENTIQDPFGTVYISKVKSQGHPA--VYRLSLNA 546
QY 911 SSEAGMNAICSSAG-CNNF-----SFTQKIQYATEFFEQSYLAIPASSWDDDFDML 961
DB 547 SLPAWQAVSPGDGEVPSFOYRAPFGNFTKCLTACMSTVGLLQAAFPSRKWM-----L 600
QY 962 TPSSC-----CRLYIS--GP 974
DB 601 TTLACDAKRWKPFDSFVATKEOQHTKLYPAQSHKPPFHGRVCMAPPGLLSSSPDGP 660
QY 975 NKDKFCPSTVNSLNCNKMGITMG-----SVRPSVQGFHKYLPWF-----LN- 1017
DB 661 TKGFF---FVPSKYPKARLSATFGFNCVNTGCGKPAVRPLVDGTGMVVFVIGIYNR 717
QY 1018 ----DRPNKCPKPGGLAAYSTVNLTS-----GQVL 1045
DB 718 TRQVDNHWIGDP--GSVVYDSSFDLFKEIGHLCHLCKAIAANSELVKPGGAQCLPSGYSI 775

QY 824 -----GEGLLGPFQKAYAP-----PILHWI-----TRGVVL 850
Db 674 EYEEEFVSLGDSVLSVSPGLOPASNTGSRGHLIVQLQELHHWVLSA VKSRWIV 733
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQ8- 874
Db 734 GLFVSIILISLVSPASRLRPASRAPLLFRPDNTIQVLLDLKYNLSABGISCTCSGLFQ8K 793
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFEVGPVYFVTTLGYNF 910
Db 794 PHSLQNNIRTSLEKKRGSGVPWASRPEATLQDPFGTVVISKVSQHPA--VYRLSINA 851
QY 911 S8EAGNNAICSSAG--CNPF-----SFTQIKYATEPPEOSYLAIPASSWDDFDWL 961
Db 852 SLPAPQAVSPGDGPEFSQVYRAPFGNFKKLTAQMTGVLQAASPSRKWM-----L 905
QY 962 TPSSC-----CRLYIS--CP 974
Db 906 TTLACDAKRGKPFDFSYVATKEQQHTRKLYFAQSHKPPFHGRVCMWAPPGCLLSSPDGP 965
QY 975 NKDKFCPTSVNSLNCNKMSITMG-----SVRPSVEQPHKYLPMF-----LN- 1017
Db 966 TKGFF---FVPSKVPKARLSATFGNCPVNTCGKPAVRPLVDTGAMVVFVGLIGVNR 1022
QY 1018 ---DRPNIKCPKGLAAYSTSVNLTS-----GQVL 1045
Db 1023 TRQVDNHVIGDP--GSWYDVSFDLPKEIGHLCHLCKATAANSSELVKPGGAQCLPSGYSI 1080
QY 1046 ASRFMAYHKLNSQD-----YTEALRAARELANITADLRKVPCTDPPEVP-P 1094
Db 1081 SSFLQMLHPECKELPBNLLPQOLSHGAVGVRQWISMAFESTTYKGS--SPQTSYSD 1139
QY 1095 YTITNVFYQYITLPEG-----LFM-----LSLCL--VP 1122
Db 1140 YLRWESFLQQLQALPEGSVLRGPTCEHWQIFMEIYGVQSALCGLVLSLICVAVA 1199
QY 1123 TPVASCLLGLDLRGLGLNLLSIYMLVDTVGFMALWDISYNAVSLINIVSAGMSVERV 1182
Db 1200 VETTHILL---LLPVLSILGIVCLV--TIMWGSWEM--GAVERAISLSILVSSVDYC 1252
QY 1183 SHITSPAIS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI-----1224
Db 1253 VHLVGYLLAGNLPHPQAEADARTQW--RTLEAVRHVGVAIVSSALTVIATVPLEFC 1310
QY 1225 LVGLGLAKAQLIQIFPFRNLITLGLHGLVFLPVLSYVGDVNPALALEQKRAE--E 1282
Db 1311 IIAFPK-----FGKIVALMTGVSILYTLTVSTALLGIMAPS-----SFRTRTSFLK 1358
QY 1283 AVAAVMVA 1290
Db 1359 ALGAVILA 1366

RESULT 6
US-11-096-568A-6778
; Sequence 6778, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; PRIOR FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6778
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(127)

; OTHER INFORMATION: Ceres Seq. ID no. 15168324
US-11-096-568A-6778

Query Match 2.8%; Score 191; DB 7; Length 127;
Best Local Similarity 44.3%; Pred. No. 6.4e-08;
Matches 35; Conservative 21; Mismatches 23; Indels 0; Gaps 0;
QY 1199 ERAKEATISMGSAVFAGVAMTNLPGILVGLAKAQLIQIFPFRNLITLGLHGLVFL 1258
Db 34 QRAKTALCTMGASVFSGITLTKLVGLVLCFSTSQIFVYVYFQMYLALVGLHGLVFL 93
QY 1259 PVLSYVGDVNPALALEQ 1277
Db 94 PVVLSLFGPPLRYTVIKEQ 112

RESULT 7
US-11-096-568A-6779
; Sequence 6779, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; PRIOR FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6779
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(85)
; OTHER INFORMATION: Ceres Seq. ID no. 15168325
US-11-096-568A-6779

Query Match 2.5%; Score 173; DB 7; Length 85;
Best Local Similarity 44.3%; Pred. No. 9.6e-07;
Matches 31; Conservative 19; Mismatches 20; Indels 0; Gaps 0;
QY 1208 MGSAYFAGVAMTNLPGILVGLAKAQLIQIFPFRNLITLGLHGLVFLPVLSYVGP 1267
Db 1 MGASVFSGITLTKLVGLVLCFSTSQIFVYVYFQMYLALVGLHGLVFLPVLSLFGP 60
QY 1268 DVNPALALEQ 1277
Db 61 PLRYTVIKEQ 70

RESULT 8
US-11-242-459-35
; Sequence 35, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 35
; LENGTH: 445
; TYPE: PRT

793 ALLSLDSKQEQASRLDVC--CCVKQOELPPPG 822
309 GLWSLYLAPLESSCQTSCHQNCSTRKTSLHFFG 340

793 ALLSLDSKQEQASRLDVC--CCVKQOELPPPG 822
309 GLWSLYLAPLESSCQTSCHQNCSTRKTSLHFFG 340

RESULT 10
US-10-392-234A-18
; Sequence 18, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 18
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-392-234A-18

Query Match 2.0%; Score 141.5; DB 6; Length 1048;
Best Local Similarity 18.1%; Pred. No. 0.011;
Matches 145; Conservative 110; Mismatches 243; Indels 303; Gaps 32;
QY 624 EINTTAEDLPFATSYIVIFLYISLALGSSYSSWRVVDKATL--GLGGVAVVLG--A 679
DB 339 EVVKTLE-----AIIIVFLVMYFLQNF-----RATLIPIAVPVLLGTFA 381
QY 680 VMAAMGPFSSYLGISSVLIVQVVPFLVLSVG--ADNIFIVLEYQLPRRPGEPREVHIG 737
DB 382 VLAAGFP-----SINTLTMFG--WVLAIGLVDDAIVVENVERVMTTEGLPPKEATR 432
QY 738 RALGR-----VAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAVLIDLLQNSAFV 792
DB 433 KSMGQIQGALVGIAMVLSAVFIPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALITLP 490
QY 793 ALLSLDSKQEQASRLDVC--CCVKQOELPPPGQEGLLGFGKAYAPFLHVI----- 844
DB 491 ALCA-----TMLAPVAKGDHGEKGFHFNRLFDKSTHYTDSVGNILR 536
QY 845 TRGVVLLFLALF-----GVSLYSMCHISVGLDQELALPKDSYLL 884
DB 537 STGRYLLLYLIIIVGWAYLVFVRLPSSFLPDEQGVFL--TMVQLPAGATQERTOKVLDEVT 595
QY 885 DYFL-----FLNR----- 892
DB 596 DYVLNKEANVESVFAVNGFGFAGRGQNTGIAFVSLKDWADRRGKKNVEAITQRTATAAF 655
QY 893 -----YFEVGAPVYFTTILGVNFSSEAGNNAICSSAGCNFSFTQ----- 932
DB 656 SQIKDAMVFAFNLPAIVELGTATGDFE-----LIDQAGLGHKLTQARNQLFGVAK 708
QY 933 -----KIQVATEPFEQSYLAIPAS-----SWDDFDIDWLTLP 963
DB 709 YPDLLVGVVRPNGLDTPQFKIDIDQEAQALGVISIDINTTLGAAMGGSYVNDPID---R 765
QY 964 SSCCRLYISGPNKDFCPSTVNSLNCNCKMCSITMGSVRPSVQFHKYLPWELNDRPNIK 1023
DB 766 GRVKVYVWSEAKYRMLPDDIND-----WYVSGSQGM 798
QY 1024 CPKGLAAVST-----SVNLTSDGOVLSRFRMAYHKPKLNKSNQDYTEALRAARELANIT 1077
DB 799 VP----FSAFSSRWYEGSPRLRYNGLPSMEILLQOAPGKST---GEAMAMBELAKLP 852

Query Match 2.0%; Score 143; DB 7; Length 456;
Best Local Similarity 25.4%; Pred. No. 0.0025;
Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;
QY 576 LNNYPAGDPRLAQAQKWEAEFLERAFQRRMAGMFQVTFATRSLED----- 623
DB 100 LPNYSVDDR-----WEE-----QRAKFSFVTVVAMLAQSTSKVQVLYGGTD 144
QY 624 -----EINRTTAED--LPFATSYIVIFLYISLALGSSYSSWRVVDKATLGLG--VAVV 676
DB 145 LFDYEVRTFNNDMLLAFISSCIAALVYILTSCVFLSFFGI-----ASIGLSCLVALF 199
QY 677 LGAVMAAGPFSSYLGISSVLIVQVVPFLVLSVGADNIFIVLEYQLPR--RPGEPREVH 735
DB 200 LYHVVFGL---QYLGILNG-----VAAFVIVIGVDDVFFVINTYRQATHLEDQLRMH 251
QY 736 IGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAVI---LDLFLQMSAFV 792
DB 252 TVQTAGK---ATFTSLTTAAAYAAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308
QY 793 ALLSLDSKQEQASRLDVC--CCVKQOELPPPG 822
DB 309 GLWSLYLAPLESSCQTSCHQNCSTRKTSLHFFG 340

RESULT 9
US-11-242-459-34
; Sequence 34, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: KITA, Shunbo
; APPLICANT: TANIMAWA, Yoshio
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703US0P
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 34
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-34

Query Match 2.1%; Score 143; DB 7; Length 456;
Best Local Similarity 25.4%; Pred. No. 0.0025;
Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;
QY 576 LNNYPAGDPRLAQAQKWEAEFLERAFQRRMAGMFQVTFATRSLED----- 623
DB 100 LPNYSVDDR-----WEE-----QRAKFSFVTVVAMLAQSTSKVQVLYGGTD 144
QY 624 -----EINRTTAED--LPFATSYIVIFLYISLALGSSYSSWRVVDKATLGLG--VAVV 676
DB 145 LFDYEVRTFNNDMLLAFISSCIAALVYILTSCVFLSFFGI-----ASIGLSCLVALF 199
QY 677 LGAVMAAGPFSSYLGISSVLIVQVVPFLVLSVGADNIFIVLEYQLPR--RPGEPREVH 735
DB 200 LYHVVFGL---QYLGILNG-----VAAFVIVIGVDDVFFVINTYRQATHLEDQLRMH 251
QY 736 IGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSLGLAVI---LDLFLQMSAFV 792
DB 252 TVQTAGK---ATFTSLTTAAAYAAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308

QY 1078 ADLRKVPCTDPAEPVFTYITNVEYEQYLT-----ILPGLFMLSCL----- 1120
Db 853 SGI-----GYDWTGMSYQERLSCNQAPALYALISLVVFLCLAALYESWSIPF 899
QY 1121 -----VPTFAVSCILLGLDLRSGLNLNLSVMILVDTVGFMALWDISYNAVSLIN----- 1170
Db 900 SVMVLVPLGVIGALLAA--TFRGLTNDVYFQVGLLTIGLSA-----KNAILIVEFAKOL 952
QY 1171 -----LVSAGVMSVEFVSHITRSPALSTKPTWLERAKEATISMSGSAVPAGVAMT 1219
Db 953 MDKEGKGLVEAMLEAVMRRLPILMTSLAFMLGVMP-----LVISSG-----AGSGAQ 1000
QY 1220 NLPGILVL-GLAKAQLIQIF 1239
Db 1001 NAVGTGLGVMTATVLAIF 1021

RESULT 11
US-10-392-234A-16
; Sequence 16, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 1046
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-392-234A-16

Query Match 2.0%; Score 140.5; DB 6; Length 1046;
Best Local Similarity 20.5%; Pred. No. 0.013;
Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;

QY 506 NOTLMGOT--SQVDWKDHFYLCANAPLTFKDGFTALALSCWADYGAPVFPFLAIGYKGD 563
Db 231 NATIICKTRLOTAEQENILLKVN-----PDGSOVLKVDVVG-----IG----GQD 274
QY 564 YSABEALIMTFSLNYP-----GDPRLAQAKLWEE--APLEEMRAFORMAGMFQ 612
Db 275 YS-----INAQNGSPASGAIKATGANALDTAKAIRQTIANLEPFPQGMKVVPYD 328
QY 613 VFTAERSLEDEINRTAEDLPATSYIVIFLYISLALGSYSWSRWVDSKATL--GL 670
Db 329 TTPVWSASIH-EVVKLTGE-----AIIIVFLVWYLFQNF-----RATLIPTI 370
QY 671 GGVAIVLG--AVMAAGFFSYLGIRSSVLTLQVVPFLVLSVG--ADNIFVLEYORLPR 726
Db 371 AVPVVLGTFGVLAAGF-----SINTLTNFG-MVLAIGLLVDDAIVVENVERVMA 421
QY 727 RFG-EPREV---HIGRALGR-VAPSMLLCSLSAICFFLGCALTPMPAPVATFALTSLAVI 781
Db 422 EBLGSPREARKSMGQIQALGVIANVLSAVFLPMAFFGSGTGI--YQFSITIVSAMA 479
QY 782 LDFLLQMSAFVALLSDSKRQERASLDVCC-----VKPQELPPPGQEGEGLLLFFOKAYAP 838
Db 480 L-----SVIVALI-----LTPALCATMLKPIEKGDHGEKGFQGFENRMFLS 522
QY 839 FULLHWITRGV-----LILLFLALFGVLSYMSCHISVGL-----DQ-----ELALPKD 880
Db 523 -TTHGYERGVASILKHRAPYLLIYVIVAGMIWMFTRIPTAFLPDBDQGVFAQVOTPPG 581

QY 881 S-----YLLDYFLFLNRYFEVGAAPVYFVTTILGYNFSSE----- 913
Db 582 SSAERTQVVVDMSREYLLER-----ESSSVSSVFTVTGTFAGRGQSSGMAFIMLKP 633
QY 914 -----AGMANI-----CSSAGCNNEFSTQKIQYATEFPEQSYLAIPASSWVDDPD----- 959
Db 634 WEERPGGNSVFLAKRAQMHFFSPKDMVFA--FAPPSVLELGNATGFLFLQDQAGVG 691
QY 960 -----WLTPSCCRLYISGPNKDKFCPSTVN-----SLNCLKNQMSI 996
Db 692 HEVLLQARNKFLMAAQNAPALQVRPNMGMSDEPOYKLEIDDEKASALGVSLADINSTVSI 751
QY 997 TMGSV-----RPSVEQFHKYLPWFINDRPN-----IKCPKGLA---AYST 1034
Db 752 AMGSSVYNDPDRGRVKRV-----YLOGRPDARMNPDDLKSWVVRNDKGEWVPFNAPAT 805
QY 1035 -----SVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDP 1088
Db 806 GRWEYSGPKLERYNGVPAMEILGEPAPLSSGD---AMAAVEE---IVKQLPKGVG--- 855
QY 1089 APEVFPYTTITNVEYEQYLT-----ILPGLFMLSCL-----LVPTPAVS 1127
Db 856 -----YSWTGLSYERLSGSQAPALYALISLVVFLCLAALYESWSIPFSVMLVPLGVI 909
QY 1128 CLLGLDLRSLNLNLSIVMILVDTVGFMALWDISYNAVSLINL-----V 1172
Db 910 GALLATSMR-GLSNDVFFQVGLLTIGLSA-----KNAILIVEFAKELHEQKGIVEAAI 963
QY 1173 SAVGMSVEFVSHITRSPALSTKPTWLERAKEATISMSGSAVPAGVAMTNLPGILVL-GLAK 1231
Db 964 EACRMELRPIVMTSLAFILGVVP-----LAISTG-----AGSGSQAHTGTVIGWMT 1011
QY 1232 AOLIQIFFERL 1242
Db 1012 ATVLAIFWVPL 1022

RESULT 12
US-10-467-657-2648
; Sequence 2648, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2648
; LENGTH: 1067
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2648

Query Match 1.9%; Score 134; DB 6; Length 1067;
Best Local Similarity 19.4%; Pred. No. 0.043;
Matches 185; Conservative 139; Mismatches 354; Indels 274; Gaps 44;

QY 452 RHLQWV-SPEAQR--NISLQIDCYAPLNPDPNTSLYDCCINSLLQVFPQNNRTLLLTANOT 508
Db 180 RAMRIWVDPKKLQNYNLSPADVGA-LSAQNQISAGSIGSLPAV--RQGT---VTATVT 233
QY 509 LMGTQSDVDKDHFLYLCANAPLTFKDGFTALALSCWADYGAPVFPFLAIGYKGYSEAE 568
Db 234 AQOGLGTAEFEFGNVLIRANT-----DGSNIYLDKVAKVGLGM-----EDYSSST 277

Qy 269 POALDSTYLGOMPGSLVLIILCSVFAVWTLVGRFVAPARDKXWVDPKK-----GTS 324
Db 177 LQPFSS-----AAMRV-----WIDPKLVGFLS 201
Qy 325 LSDKLSPSTHLLGQFFQGWGTWASWPLTILVLSVIPVAL--AAGLVFTELTDPVEL 382
Db 202 IDD-----VSNIRGNVQ-----VPAGAFGAPGSSAQELTATLAVK 239
Qy 383 WSAPNSQARSEKAFHDOHFGFPFRNTQVILTAPNRSSYRSDSLGLGPKNFGLDLDL 442
Db 240 GTLDPQFGQVVLPRANEDGSLVRLADVARLELKGESYNISRLNGTPTVGAIGQLSPCA 299
Qy 443 ELLE-----LOELRLHQSPEAQRNISLQDICYAPLNDPNTSLY-----DCCINSLLQ- 492
Db 300 NAIQTATLVKORLAELSAPFPE-----DMQYSV--PYDTSREVDVAIEKVHTLEA 349
Qy 493 -----YFQNRRTLL--LTANQTLMGTSQVDMKDHFLYCANAPLTFKDTALAL 540
Db 350 MVLVFLVLMPLFLQVRYTLIPSIVVPVCLLGLTMYM-----YLLGFSVNMMTMFGVLA 404
Qy 541 SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSILNNYPAGDPRLAQAKLWEEAPLEEM 600
Db 405 GILVD-----DAIVVENVERIMA--EGISPA-----EATVKAM 437
Qy 601 RAFORMMAGMFQVTFTAERSLEDEINRTAEDLPI--PATSYIVIFLYLSIAL----- 651
Db 438 KQVSGAIVGITLV-----LSAVFLPLAFMAGSVGVYQOFSVLAVSILFS 483
Qy 652 -----GSYSMSR-----VWVDSKATLGLGGVAV 675
Db 484 GFALTFTPALCATLKPTEGHEKRGFFGAFNGFARVTRYSLNLSKLVARAGRFLM 543
Qy 676 VLGAVMAAGPFSYLGIRSLVILQVFPFLVLSVGADNIFIFVLEYQRLPRPGPREVH 735
Db 544 VTAGLVAMLGYP-YLRLPFAVPAEDLGYMVVDV-----QLP--PGASR--- 584
Qy 736 IGRALGRVAPSMLLCSLSBAICFFLGALTPMPAVRTFA-----LTSGLAVIDFLQ 787
Db 585 -----VRTDATGEELERFLKSRVAVSVFLIS 611
Qy 788 MSAPVALLSLDSKQASRLDVCCVKPQELPPGQEGLLGCFQKAYAPFLLHWITRG 847
Db 612 GFSF-----SGQDGAALAF-----PTFKWSERG 636
Qy 848 VVLLFLALFGVSLYSWCHISVGLDQELALPKDSYLL----- 884
Db 637 -----AEQAAAEIAALNEHFAIPDDGTWAVSPPPINGLNSGGFALRLMDR 684
Qy 885 -----DYFL-----FLNRVFE--VGAPVYFV-----TTLGYNFSSEAG 915
Db 685 SGVGREALLQARDTLLEIGIQTNPKFLYAMMEGLAEAPQLRLLDREKARALGVSFETISG 744
Qy 916 -MNAICSSAGCNFFSTQKIYATEPPEQSYLAI PASSWVDDFDIMLTPSSCCRLYISGP 974
Db 745 TLSAAGFSEVINDFTNAGROORVVIQAEQG-----NRMTPESVLELV--- 787
Qy 975 NKDKFCPSTVNSLNCNLCMSITW--GSVRPSVEQHKYLPWFILNDRPNIKCPKGLAAY 1032
Db 788 -----PNAAGNLVPLSAFYSVKWEGPV-----QLVRY-----NGYPSIRIVGDAAPGF 831
Qy 1033 STSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAFEV 1092
Db 832 ST-----GEAMA-----EMERLASQLPAGI----- 851
Qy 1093 PPYITITNVFEQYITI-LPBEGLFMLSICLVPTFAVSCLLGLDLSGLNLLSIVMLT-- 1149
Db 852 -GYEWTLGSLYQEKVSAGQATSLFALATLV-----FLLLVALYESWSIPLSVMLIVPI 903
Qy 1150 --VDVTGFMALWDISYNAVSLINIVSAGMS-----VBFVSHITRSPAISTKPTWLE- 1199
Db 904 GAICAVLAVMVGNSNDVYFKVGLITIIIGLSAKNAIIVF-----AKELWEQG 952
Qy 1200 -RAKEATISMSAVFAGVAMTNLP---GILVL-----GLAKAQIIQIFFRNLNLTLL 1249

Db 953 HSLDAEAIEARLFRFRPIINTSMAFILGVIPLALASGAGAASQRAIGTGVIGCMLSATFL 1012
Qy 1250 GLHLGLVFLPV 1260
Db 1013 GVL-----FVPI 1019
RESULT 17
US-11-203-806A-11
; Sequence 11, Application US/11203806A
; Publication No. US20060051737A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Silverman, Joshua
; TITLE OF INVENTION: Nucleic and Amino Acid Sequences of
; FILE OF INVENTION: Human Rifi and Methods of Use Thereof
; FILE REFERENCE: 600-1-298N
; CURRENT APPLICATION NUMBER: US/11/203,806A
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/601,405
; PRIOR FILING DATE: 2004-08-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2426
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-203-806A-11
Query Match 1.7%; Score 120.5; DB 7; Length 2426;
Best Local Similarity 19.9%; Pred. No. 1.6;
Matches 184; Conservative 106; Mismatches 294; Indels 341; Gaps 46;
Qy 11 LWALLRLAQSEPYTHIQPGYCAPFYDCGKNPELGSMLTSLNSVCLSNTP----- 62
Db 259 LWPLFVKLLGK---TLHRSG-----SPINSLQLLEELGFRSGTGMKIKAFI 302
Qy 63 ARKITGDHILLOKICPRLYGTGPTQACCSAKQLVSLSEASITKALLTRCPACSDNFVN 122
Db 303 AWKSLIDNFAL-----NPDILCSAKRLKLMQPLS-----S 333
Qy 123 LCHNTCSNQSLFNTVTRVAQLGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRVP-AAA 181
Db 334 IHVRTETALTALKEVWVWYLLMLRGP-QLPA-----NFEQVCVELIQS 374
Qy 182 TLAVGTGCVGYSGALCNAQRWLNFGQDTGNGLAPLDITPHLLEPGQAVSGGQPLNEGVA 241
Db 375 TISVDSIPSQG---NSR-----GSASPGLSP-----LTPGH---KGASPYGSPRG 415
Qy 242 RNESQGDVATCSCODCAACFCFATARPQALDSTFVLGOMPSLVLIIILCSVFAVVTL 301
Db 416 NLSSNTGG-----MAAIPSI-----QLLG-----LEMLLHF 441
Qy 302 LVGFRVAPARDKSKW---DPKKGTSLSDKLSFS---THTLLGQFFQGWGTWASWPLTIL 356
Db 442 LLQPEVLFAKQKQKIVLSLEPLEHPLISSPFSKYAHTL-----ITAV 485
Qy 357 VLSVIPVVALAAGLVFTELTDPVEL--WSAPNSQARSEKAFHDOHFGFPFRNTQVILTAP 415
Db 486 HDSFVSVGKDASDAVVSAIWKELISLVKSVTEAGNRKEK-----SGSEVLTL- 532
Qy 416 NRSSYRYSLLGPKNF--SGILDLDLLELLELQER-LRHLOWSPQEAQRNISLQDICY 472
Db 533 -----LLKSLENIVKSEVPFVSKTLVLMETVKGLPKPKVIGSPAYQ--VANMDI-- 579
Qy 473 APLNPD-----NTSLYDCCINSLLQYFQNRRTLLLTANQTLMGTSQVDMKDH 522
Db 580 --LNGTALFLIQLIIFNNILLEGVED-EKFFLNLTLV---GCVLSGTSPPLAFSDSV 632
Qy 523 LYCAN-----APLT-----FKDGTALALSCMADYGAPVFPFL 554
Db 633 LTVINQNAQLVNKEHLWRMWSMIVSPLTDVIHQTNVQGDALHNFSIAYGALTLPIN 692

Query Match 1.7%; Score 115.5; DB 7; Length 914;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 208; Conservative 140; Mismatches 381; Indels 375; Gaps 55;
Qy 55 VSCLSNTPARKITG-----DHLIL-----LQICPRLYTGP-NTQACCSAKQLVS 98
Db 19 VICLRPP--KVLGYRLEPPLTLACTLEGMYNIEELERDINGTLLSQATCELCD--G 74
Qy 99 LEASLSITKALLTRCPACSDNFVNLHCHNTCS-PN---QSLFINVT-----RVAQLGAG 148
Db 75 NENFMVWVNLGDCRCVCEPTFVNTSRSCASEPNILTGGLCFSGTGNFRLRISAARYG 134
Qy 149 QLPVAVYAEFYQHSFABQSYDSCSRVRVPAATLAVGTM-----GVYGSALCNAORWLN 204
Db 135 EVGSLTSEWFAKYL---QSSAAACWVYANLTSCQALGSMCMNMNSYDFATFDACGLFQ 191
Qy 205 FQDGTGNGIAPLDTIFHLE-----PGQAVSGIOPLNEGVARCNEQSGDDVATCSCQD 258
Db 192 FIFENTAGLS---TVHSISFWRQMLPWFVGDQL-----GLA----- 225
Qy 259 CAASCFAIARPAOLDSTFYLGQMPGSLVLIILCSVFVAVTILLVGFVRVAPARDKSKVD 318
Db 226 -----FQVLSST-----SLPTNFSF----- 240
Qy 319 PKGFTSLSDKLSF--STHTLQFPQGWGTWVASHPLTILVLSVIPVVALAAGLVFTLT 376
Db 241 --KGEDQNTKLFVAASYDIRGNFLK-WQT-----LEGV----- 272
Qy 377 TDPVELMSAPNSQASEKAFDQHFQPPPTN-----QVILTAPNRSSRYRDSLLGPK 430
Db 273 -----LQCPDTETRLNAY---SGFTYQOCEIPISKILIDFPPIFY----- 314
Qy 431 NFGSILDLDLLELLELQERLRHLQVWSPQARNISLQICVAPLNPDTSLYDCCINSL 490
Db 315 -----DVLEYTDENQHVILAV---PVLNMLNH-----NKI 344
Qy 491 LQYFQNNRLLLTANQTLMGQTSQYDWDKHELYCANAPLTKDGTALALSCMA---DYG 547
Db 345 FVNQDSNGKWLITRIFLV---DAVSGRENDL---GTQPRVIRVATQISLHVLPNTIN 399
Qy 548 APVFPFLAIGGKGYKDYSAE-----ALIMTFSLNNYPAG-DPRLAQAKLWEAFLEEM 600
Db 400 GNYPPLITITAYSDIDIKNANSQVSVSVTYEMDGEAHVQTDIALGVGLGLAVLASL 459
Qy 601 --RAFORMAGMPQVTTAERSLEDRINTTADLPFIPATSYIVIFLYISLALGSSWS 658
Db 460 LKTAGWKIRIGSPWIDL-----QTVWKFVLYVYAGDLANVFPIITVGTGLY--W- 505
Qy 659 RMVVDKATLGLGGVAVVLGAVMAAGFPFSLYGLRSLVLTQVVPFLVLSVGDANIFIV 718
Db 506 --LIFFAQ---KSVSVLLPMPQBERFVTVGCAFKALQFLHKLISQITID---VFF 557
Qy 719 LEYORLPRRPGEPREHVHIGRALGRVAPSMLLCSISEACFFLGTALTPMAVRTF----- 772
Db 558 IDMER-----PK-----GKVLKAV-----EGEGGVSRSATVPVSIWRTYFVANEW 596
Qy 773 -----ALTSGLAVIDFLQMSAFVALLSLS-----KRQEASRLDCCCCKPQEL 818

TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 2923
LENGTH: 914
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-2923

Qy 738 RALGR-----VAPSMLLCSLSEACIFFLGALTMPAVRTFALTSGLAVILDFLLQMSAFV 792
Db 433 KSMQIQOALGIANVLAVFVNPVAFPGSGSGAI--YRQFSITIVSAMLSVLVALITP 490
Qy 793 ALLSLSKROBASRLDVCCCVKQPQLPPPGGEGLLIGFPQKAVAPFLLHWI----- 844
Db 491 ALCA-----TMLKPIAKDGHGSGKGFGEFGRNFMFEKSTHHTYDVSVGILR 536
Qy 845 TRGVULLFLALFGVSLY-----SMCHISVGLDQE----- 874
Db 537 STGRYLYLLIIVGMAFLFVRLSSFLPDDQGVFMVWQLPAGATQERTQKVLNEVTH 596
Qy 875 --LALPKDS-----YLLDYFLNLR----- 892
Db 597 YLITKEKNVSVFAVNGFAGRGQNTGIAPVSLKQWADRPGEENKVEAITMRATRAFS 656
Qy 893 -----YFEVGAPVYFVTTLGYNFS--SEAGM-----NAICSSAGCANNFSPT-- 931
Db 657 QIKDAMVAFNLPALVELGTATGDFDELIDQAGLGEKLTQARNQLLAENAKHPDMLTSV 716
Qy 932 -----QKIQVATEPEQSYLAIPAS-----SWDDDFIDWLTSPSSCCRLYI 971
Db 717 RPNGLDTPQFKIDIDQKQAQALGVSINDINTLGAANGSGSVNDFID---RGRVKKYV 773
Qy 972 SGPNDKFCPSTVNSLNCNKMSITMGSVRPSVEQFHKLFPWFLNDRPNIKCPKGLAA 1031
Db 774 MSEAKRMPLDDIGD-----WYVRAADGGQVMP--FSA 803
Qy 1032 YST-----SVNLTSBGQVLASRFMAYHKPLKNSQDYETALRAARELANITADLRKVP 1085
Db 804 FSSSRWEYSGPRLERVNGLPSEIILQQAAPGKST---GEAMELMEQLASKLPTGV----- 855
Qy 1086 TDPAEVFYPTTINVEYQYL-----TILPGLFMLSCL-----VPT 1123
Db 856 -----GYDTMGMSQBERLSGNQAPSLYALSIVVFLCLALAYESWISIPSVMLVPL 907
Qy 1124 FAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAWMDISYNAVSLNL----- 1171
Db 908 GVIGALLAA--TPRGLTNDVYFQVGLTITIGLSA-----KNAILIVEFAKDLMDKEGKL 960
Qy 1172 ---VSAGVMSVEFVSHITRSPASTKYPTWLERAKEATISMSGAVPAGVAMTNLPGLVYL 1227
Db 961 IEATLDVAVRMRLRPILMTSLAFLGVMP-----LVISTG---AGSGAQNAGVTGVM 1008
Qy 1228 -GLAKAQLIOIFF 1239
Db 1009 GGMVTATVLAIFF 1021

RESULT 20
US-11-072-512-2923
Sequence 2923, Application US/11072512
Publication No. US20060029945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOTUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO

```
Db 597 NEIQTVRKINSLFQVLTVL--PFEVVGFKNLALMDSSSLRNPFSYIAPYSCI-----649
QY 819 PPGQCEGLLLGFQKAYAPFLHWTGRGVVLLFLALP-----GVSLYMCHI 867
Db 650 -----IRYAVSAALWLAIGIIQVFFAVFERIEDKIQFVDCLSMSNI 694
QY 868 SVGLDQELALPKDSYLLDYFLFNRYFEVGAPVY-----FVVTLYGNFSSEAGMNAICSS 922
Db 695 SV-----FLLSHKCF--GYVIHGRSVGHADTNMEEMNNLKKREA--ENLCSQ 738
QY 923 AGC--NNFSFTQKIQVATEPPEQSYLAIPASSWVDWDFIDWLTSSCCRLYISPNKDKFC 980
Db 739 RGLVPNTDQGTPEIAISNQW--RQHYDRIHET-----LIRKNGPARLLSS--782
QY 981 PSTVNSLNCNKMSITMSVRSVEQFKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSS 1040
Db 783 ASTFEQSIKAYHNMKFLGSP---IDHVHKEMDYFIKDKLLLE-----822
QY 1041 DGOVLASRFMAYHKPLKNSQDYTE 1064
Db 823 --RILGWEMF---EPMEKSIIFYND 841

RESULT 21
US-10-392-234A-30
; Sequence 30, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acRAB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-392-234A-30

Query Match 1.6%; Score 113; DB 6; Length 1034;
Best Local Similarity 18.8%; Pred. No. 1.9;
Matches 161; Conservative 115; Mismatches 270; Indels 312; Gaps 38;

QY 576 LNNYPA-----GDPRLAQKLWEAFLEEMRAFORMAGMFQVTTAERSLE-DEI 625
Db 281 INKPAAGIGIKATGANALDTAKAKLAELQPFPPQMKVLYPYDTTPFVQLSIHEV 340
QY 626 NRTTADLPFATSYIVIFLYLSLALGYSYSSRVNVDGKATL--GLGGVAVVLG--AVM 681
Db 341 VKTLFPAIMLV---FLVWVLFQ-----NRRATLIPTIAPVVLGTFALL 383
QY 682 AAMGPFYSLGIRSLVILQVPELVLSVG--ADNIFIVLEYOR-----LPRRGPPEPR 733
Db 384 AAFGY-----SINTLWFG--WVLAIGLLVDDAIVVENVERVNMEDKLP--PKEATE 432
QY 734 VHIGRALGR--VAPSMCLLSLSEACFLGALTMPAPVTFALTSLGLAVILDFLLQMSAFV 792
Db 433 KMSQIQGALVGAMVLSAVFIPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALILTP 490
QY 793 ALLSLDSKQEASRLDCCVCKPQELPPPCQGGGLLGFQKAYAPFLHWM-----843
Db 491 ALCATLLKPVSAEH-----HENKGGFGWNTTFDHSVNHVNTSVSKILG 535
QY 844 -----ITRGVVLILLFLAL-----FGVSLYSMCHISVGLDQELALPKDSYL 883
```

```
Db 536 STGRYLLIYALIVAGNV--VLFRLPSSFLPEEDQGVFL-TMIQLPAGATQERTQKVLQDV 593
QY 884 LDYFLFLNRYFEVGAPVYFVTTL--GYNFSSEAGMNA-----918
Db 594 TDYLLKNEK-----ANVESVFTVNGPFSFGQA--ONAGMAPVSLKPMWEERNGDENSABAVI 647
QY 919 -----ICSSAGCANNFSFTQ-----932
Db 648 HRAKMELGKIROGFVFPNMPAIVELGTATGDPDFELIDQAGLGHDLALTOARNQLLGWAAQ 707
QY 933 -----KIQVATEPPEQSYLAIPAS-----SWVDDFIDWLTTP 963
Db 708 HPASLVSVRPNGLEDTAQFKLEVDAQKALGVSLSDINQITISTALGGTYVNDPID--R 764
QY 964 SSCCRLLYISGPNKDKFCPSTVNSLNCNKMSITMSVRSVEQFKYLPWFLNDRPNIK 1023
Db 765 GRVKUYQADAKFRMLPEDV-----DKUYVR 791
QY 1024 CPRGGLAAYSTSVNLTSDQVILASRFMAYHKPLKNSQDYTEALRA-----ARELAANITA 1078
Db 792 SANGENVPFSA---FTTSHWVYGSPLERYNGLPSEIQCERAAAPGTTSSGDAMALMENLAS 848
QY 1079 DLRKVPDTPAPEFVFPYTTITNVFYEQYLT-----ILPEGLFMLSCLVPTF-----AVS 1127
Db 849 KLPAGIG-----YDWTGMSYQERLSGNOAPALVAISFVVVFLCLAALYESWSIPVS 899
QY 1128 CLL-----LGLDLRSGLNLLSIYMLV---DTVGFMAWDISYNAVSLINL-----1171
Db 900 VMLVPLGIVGVLLAATLNFQKNDVYFMVGLLTTIGLSA-----KNAILIVEPAKOLMEK 954
QY 1172 -----VSAVGMSVRFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLP 1222
Db 955 EKGGVVEATLMAVRMLRILMTSLAIFLGVL-----LAISNG-----AGSGAQNVA 1002
QY 1223 GILVL-GLAKAQLIQIFP 1239
Db 1003 GIGVMGVMVSATLLAIFP 1020

RESULT 22
US-11-082-389-362
; Sequence 362, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
```


QY 1127 SCILLGLDLSGLNLSIVMILV-DTVGPMAL-WDISYNAVSLINLVSAGVMSVE--FV 1182
Db 432 -----TIVMLLIWQTNIFLALCFPLIFGVSVEIYLLAVLTILKLEGGWV 474
QY 1183 SHITRSFAISTKPTW-----LERAKEA--TISM-----GSAVFAVAMTNLPGILVLGL 1229
Db 475 PLVFATPFLTVIMYINWYGVSKYQSEVRERISWDFMRELGST-----LGTIRIRGI--GL 527
QY 1230 AKAOILQ-----IPFRLNLITLGLLHGLVFLPVLISVYVGPDPNP 1271
Db 528 LYNELVQGISIFG---QFLTLPAIHSTIIF--VCIKYVPVPVP 568

RESULT 24

US-11-096-568A-30411
; Sequence 30411, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30411
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(772)
; OTHER INFORMATION: Ceres Seq. ID no. 4958917
US-11-096-568A-30411

Query Match 1.6%; Score 109.5; DB 7; Length 772;
Best Local Similarity 20.5%; Pred. No. 2.3;
Matches 145; Conservative 89; Mismatches 211; Indels 261; Gaps 41;

QY 670 LGGVAVLGAVMAAMG-----PFSVLGIRS-----SLVT--LQVVP-----FL 705
Db 34 LGIAFTQLGVVYVGMGTSPLYVFSVSKVPIRSEVDVLGSLVITYITIAVPLAKYVFP 93
QY 706 VLSV-----GADNIFIFVLEYQRLPRPG--BPREVHIG-----RALG----- 741
Db 94 VLKANDNGEGGTFFALYSLICRYAKVKNLKNQPADEQISSPRLKLTPELERALGKEAL 153
QY 742 ---RVAPSMILCSLSEAICFFLGALTPMPAVRTFALTSGL-AVILDF---LLQMSAFVAL 794
Db 154 ETGYLKTLLLLVLMTGSMIIIGDGLITPAMSVMSAMSGLQGEVKGFGTNALVMSIVIL 213
QY 795 LSLDSKQREASRLDVCCCKVQELPPGQGGELLIGFQKAVAPFLHWTGVLVLLFL 854
Db 214 VALFSIQRF-----GTGK---VGFL---PAPVLALW-----FF 240
QY 855 ALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLPLNR-----YFVCGAPVYFVT----- 904
Db 241 SIGALGIYNLLKYDFTVIRAL---NPFYIV---LFFNKNKSQMSALGGCVLCITGABAM 294
QY 905 ---TLGYNFSSEAGMNAICSSAGCNNSFTQKIQYATEPPEQS-----YLAIPAS-SWDDF 957
Db 295 PADLGHFSVRSIQMAFTCVVPFCLLAYMGQAAYLTKEPEASARIFYDSVPKSLFPWPFV 354
QY 958 IDWLTSPSCCRLYISGPNKDKFCPTVNSLNCNKCISITMGSVRPSVEQFHK----- 1010
Db 355 IATLAAMTASQAMISA-----TFSCVQKQAMAL---GCFFRLKIHTSKRIGQ 399
QY 1011 -YLP---WFLNDRPNIKCPKGLAAYSTVNLTSQGVLASRFMAYHKPLKNSQDYTEAL 1066
Db 400 IYIPVINWFL---MIMC-----ILV-----VSIF 420
QY 1067 RAARELAANATADLRKVPDTPAPEVFPYTTINVFYEQYLTILPGLFMLSICLVPTFAV 1126

Db 421 RSTTHIAN-----AYGIAEV-----GVMVSTVLV----- 445
QY 1127 SCILLGLDLSGLNLSIVMILV-DTVGPMAL-WDISYNAVSLINLVSAGVMSVE--FV 1182
Db 446 -----TIVMLLIWQTNIFLALCFPLIFGVSVEIYLLAVLTILKLEGGWV 488
QY 1183 SHITRSFAISTKPTW-----LERAKEA--TISM-----GSAVFAVAMTNLPGILVLGL 1229
Db 489 PLVFATPFLTVIMYINWYGVSKYQSEVRERISWDFMRELGST-----LGTIRIRGI--GL 541
QY 1230 AKAOILQ-----IPFRLNLITLGLLHGLVFLPVLISVYVGPDPNP 1271
Db 542 LYNELVQGISIFG---QFLTLPAIHSTIIF--VCIKYVPVPVP 582

RESULT 25

US-11-096-568A-30410
; Sequence 30410, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30410
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(827)
; OTHER INFORMATION: Ceres Seq. ID no. 4958916
US-11-096-568A-30410

Query Match 1.6%; Score 109.5; DB 7; Length 827;
Best Local Similarity 20.5%; Pred. No. 2.6;
Matches 145; Conservative 89; Mismatches 211; Indels 261; Gaps 41;

QY 670 LGGVAVLGAVMAAMG-----PFSVLGIRS-----SLVT--LQVVP-----FL 705
Db 89 LGIAFTQLGVVYVGMGTSPLYVFSVSKVPIRSEVDVLGSLVITYITIAVPLAKYVFP 148
QY 706 VLSV-----GADNIFIFVLEYQRLPRPG--BPREVHIG-----RALG----- 741
Db 149 VLKANDNGEGGTFFALYSLICRYAKVKNLKNQPADEQISSPRLKLTPELERALGKEAL 208
QY 742 ---RVAPSMILCSLSEAICFFLGALTPMPAVRTFALTSGL-AVILDF---LLQMSAFVAL 794
Db 209 ETGYLKTLLLLVLMTGSMIIIGDGLITPAMSVMSAMSGLQGEVKGFGTNALVMSIVIL 268
QY 795 LSLDSKQREASRLDVCCCKVQELPPGQGGELLIGFQKAVAPFLHWTGVLVLLFL 854
Db 269 VALFSIQRF-----GTGK---VGFL---PAPVLALW-----FF 295
QY 855 ALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLPLNR-----YFVCGAPVYFVT----- 904
Db 296 SIGALGIYNLLKYDFTVIRAL---NPFYIV---LFFNKNKSQMSALGGCVLCITGABAM 349
QY 905 ---TLGYNFSSEAGMNAICSSAGCNNSFTQKIQYATEPPEQS-----YLAIPAS-SWDDF 957
Db 350 PADLGHFSVRSIQMAFTCVVPFCLLAYMGQAAYLTKEPEASARIFYDSVPKSLFPWPFV 409
QY 958 IDWLTSPSCCRLYISGPNKDKFCPTVNSLNCNKCISITMGSVRPSVEQFHK----- 1010
Db 410 IATLAAMTASQAMISA-----TFSCVQKQAMAL---GCFFRLKIHTSKRIGQ 454
QY 1011 -YLP---WFLNDRPNIKCPKGLAAYSTVNLTSQGVLASRFMAYHKPLKNSQDYTEAL 1066
Db 455 IYIPVINWFL---MIMC-----ILV-----VSIF 475

Db 958 LRPIIMTTAAWAGLIPLLYATGAGAVRSFISIGIVI--VAGLSIGTITFTLFLVLPVVSIVA 1016
QY 1267 PDVNPALALQXR 1279
Db 1017 TEHKPLPVDPENK 1029

RESULT 28
US-11-087-099-4135
; Sequence 4135, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4135
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Sulfolobus tokodaii
US-11-087-099-4135

Query Match 1.6%; Score 108.5; DB 7; Length 648;
Best Local Similarity 19.5%; Pred. No. 2.2;
Matches 136; Conservative 99; Mismatches 249; Indels 213; Gaps 37;

QY 604 QRMAGMFOVTFATERSLEDEINRTAEDLPFATSYIVIFLYISIALGSSWSRWVD 663
Db 3 BRETGARDIGISSDRLRRSLGKFE-----LLYLSLGGHIGSGW---LFA 45
QY 664 SKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGCADNIPFVLEYQR 723
Db 46 SLATAYAGASALSIIAG-----ILVNFGLVVAEIGA-----A 81
QY 724 LPRRPGEPREVHI--GRALGRVAPSMLLCSLSAICFFLGLALPMPAVRTFALTSLGLAVI 781
Db 82 IPKSGGIGTRVPHYTHGGLVGYLI-----TWAVFLSAAS-VPAIRAAAAIEIGSY 130
QY 782 LDFELLQMSAFVALLSDSKRQASRLDVCCVQBELPPQEGE---LLGFQKAYAP 838
Db 131 YPOLITGTGF-----DGTVTI-----LTPGIGLAGLLILFPFLNYPGV 171
QY 839 FLIHMITRG-----VLLLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFN 891
Db 172 NILGKVTGAGWKKLIPTITFLALLALDLS--ANFSLGG---FLPSAEYVNG---GSS 224
QY 892 RYFEGAPVYFVTLGYNFSSEAGMNAICSSAGCENNFSFTQKIQYATEFPEQSYLAIPAS 951
Db 225 GIYFSAVLVAIPSTGVIFS-----YLFGRQAVEYGB-----257
QY 952 SWDDFIDWLTSSCCRLYISGNKDKFPCSTVNSLNCMLKMSITMGVSRSVEQ--FH 1009
Db 258 -----GKNPKKIDIPFAVIG-----SLLIAIVLYTLQVAFV 288
QY 1010 KYLPW---FLNDPN--IKCPKGLAAVSTVNLTSQG-QVLASRFMAYHKPLKNSQDYTE 1064
Db 289 GGLDMNKVLNYESGKLVVTPGNSALSTAV--TASGVPIASGPFVLVLSRLASVS-----341
QY 1065 ALRAA---RELAANITADLRKP--GTDPAFEPVFPYTTINVFY---BOYLTLPEGLFMLS 1117
Db 342 GLAAAPPTALAVLLTIDAVVSGTG---WIYGTSTRTLYAFASNGY---LPEIFLKIG 395
QY 1118 LCLVPTFA--VSCLLGLDLRSGLLNLLSLIMVLVDVTFGMALWDISYNVSLINLVSAGV 1176
Db 396 KTKIPTYSLIAALIIGF-----IFLLPPSWVALVGFITSSATVLTYYIMGGIG 442
QY 1177 MSV--EFVSHITRESFAISTKPTWLERAKEATISMGSA-----VPAGVAMTNLPG 1223
Db 443 LTVLRKHAABLRNPFKLPAAPI---TGAITATLAAGLIVVWSSFAVLFFYFTGVFL--GLEFL 498

QY 1224 ILV-----LGLAKAQLIQIFFRFLNLLITL-LGLL 1252
Db 499 FFIFYADKMLGINK--VWSIIVGVINLAINLGMGLL 532

RESULT 29
US-11-087-099-11457
; Sequence 11457, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11457
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293
US-11-087-099-11457

Query Match 1.5%; Score 106; DB 7; Length 443;
Best Local Similarity 24.3%; Pred. No. 2;
Matches 74; Conservative 43; Mismatches 102; Indels 86; Gaps 15;

QY 1076 ITADLRKVPQTPAPAFVFPYTTINVFYEQVLTLPGLFMLSCLVPTFAVSCLLGLDL 1135
Db 96 ITLDYLILP-----VITYLISASFGQSLLPAPFWWILGIAI---FNTVINLIGIKV 145
QY 1136 RSG-----LLNLLISIVMI--LVDTVGFMAWDISYNAVSLIN-----LVSAGVMSVEP 1181
Db 146 VSKFSSWLLFGLQGVLIATFIVGTIVANGTVSFNVSSFFYNASHFDLSGVVQATGIVI--203
QY 1182 VSHITRSP-AISTKPTWLERAKEATISMGSV-----FAGVAMTNLPG--1223
Db 204 VSYL--GPDIAISTLS---BEANDPTKSGIKAVLSILIGSLFVIVTFFAGFVAPYSEL 258
QY 1224 -----ILVLG--LAKAQLIQIFFRFLNLLITLGLLHGLV-----FLPVLS 1263
Db 259 NPDPTAPLTTLGKVGSWLVHLAEFTLVLSFGLAGGQGTAVSRILYMSGRDILPRLS 318
QY 1264 YVGPDPVNPALALEQKRAEBAVAAMVVASCPNHPRSRVSTADNIYVNHHSFEGSIKAGAI 1323
Db 319 YLHPRYNTEW-----VAIILVGI-----VSVALSLSLSLTTSVNSLISFGALFG 361
QY 1324 FLPN 1328
Db 362 FLSLN 366

RESULT 30
US-11-082-389-354
; Sequence 354, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-131PCPN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09

```
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 354
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-354

Query Match          1.5%; Score 105.5; DB 7; Length 783;
Best Local Similarity 20.0%; Pred. No. 4.9;
Matches 145; Conservative 108; Mismatches 268; Indels 205; Gaps 34;

QY 628 TTAEDLPFATSVIVIFLYISLALGSYSWSRVMVDSKATLGLGGVAVVVLGAVMAAMGPF 687
DB 194 TTSEIIGIGIAFLVITF-----GSLIAGLPLITAVIGVGIGALAVLAT----- 240

QY 688 SYLGIRSSVILQVFPFLVLSVADNIFIFVL-----EYQRLPRRGEPREVHIGRALGR 742
DB 241 AFTDLNVTPLAVM--IGLAVGID-YALFILSRVYAEYKMER--ADAAGMAVGTAGSA 295

QY 743 V--APSMMLCSSEATCFPLGALTTPMPAVRTFALTSGLAVIDLDFLLQMAFAVLLSLDSK 800
DB 296 VVPAGATVIAIALVALLIADIGFLTAM-----GISAAFTVFVAVLIALTFIPALLGVFG 349

QY 801 RQESRLDVCCCKVQPELPPGOG-----EGLL-----LGFQKAYAPFLH 842
DB 350 HAPKGI-----PGIGNPTKQWQEQALNRRSKRSWKLQK--APGL-- 392

QY 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPVVF 902
DB 393 -----VAVVVLGLGALTIPAM-----NLQLSLPDS----- 419

QY 903 VTILGNFSSEAGNNAICSSAGCNFSFTQIKIYATEFFEQSVLAIPASSWDDFDWLT 962
DB 420 TSNIDTTQROSADLMAEGFGAGVN-----APFLVIVDTHEVNADSTALQ 463

QY 963 PSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVQFHKYLPWFILDRPNI 1022
DB 464 P-----LIEAQEPBEGEF-----DREQARFATYM--VYQTYSN 496

QY 1023 KCPKGGIAAYSTSVNLTSQVLASRPMAYHKPL-KNSQDYTBALRAARELAANITADLR 1081
DB 497 NIDVKNAQIISVNDFFTA-AQILVT---PYTGPADKETPELMHVLRAQE-----AQIE 545

QY 1082 KVPQTD---PAFEVFPYTIINVFVEQVLTILPEGLFMLSICLVPTEFAVSLILG-----L 1133
DB 546 DVTGTGLTGTGFTAVQLDIT-----EQLEDAMP-----VYLAVVVGLAIFLLILVFRSLV 596

QY 1134 DLRSGLNLLSI-----VMILVDVTFPMALWD-----ISNVASLNLVSAVGMSV----- 1179
DB 597 PLVAGLGLSVGAAGATVLVWQEGGGFNTPGPLISFMPFLIGVTFTGLANDYQVFL 656

QY 1180 -----EFVSHITRSFAISTKPTWLERAKEATISNGS-----AVFAGVAMTNLPG 1223
DB 657 VTRMRHVTHNGKQGPQSKVTPVEQSVIEGFTQGSRVVTAALIMITAVFVAFIDQLPFP 716

QY 1224 ILVVG--LAKAQLTIQIFFRNLNLIITLGLL-HGLVFLPVILSVVGDVN-PALALQKR 1279
DB 717 IKIFGALGAGVFFDAFFIRMLVPASGFLMGKATWMPKWLDRILPSLDIEGTALKEW 776
```

```
QY 1280 ABEAVA 1285
DB 777 EEKQAA 782
```

RESULT 31

```
US-11-087-099-11792
; Sequence 11792, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11792
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Methanosarcina barkeri
US-11-087-099-11792
```

```
Query Match          1.5%; Score 105; DB 7; Length 324;
Best Local Similarity 23.2%; Pred. No. 1.5;
Matches 90; Conservative 48; Mismatches 114; Indels 136; Gaps 21;
```

```
QY 527 NAPTFKDGKTALALSCMADYG-----APVPPFLAIGGYKGYSEAEALIMTFLNNY 579
DB 22 NGYLTGTGLFILLSLIMVYGIKESARLAIFISFIEVSG-----LLIVIVT----- 67

QY 580 PAGDPRLOAKLWEAFLEEMRAFORMAGMEQVT---FTAERSLED-----EINR 627
DB 68 --GLPYLGTVNYFETPDL-----AGIFKASTLIFPAFLGPDIVRLSQETKEAEK 115

QY 628 TTAEDLPFATSVIVIFLYISLALGSYSWSRVMVDSKATLGLGG-----VAVVLG--- 678
DB 116 TTPKALLI--AIFFTVFLMVCVATVAVS-----MLDPR-VLGLSGAPLADVAISLGNKA 167

QY 679 -AVMAAMGFYSYLGIRSSILVILQVVPFLVLSVADNIFIFVLEYQRLPRRGEPREVHIG 737
DB 168 FVLSMWIALFSTMN-----TVLVMLGGSRIYVMADSGALP----- 204

QY 738 RALGRVAPSMMLCSLSEALCFPLGALTTPMPAVRTFALTSGLA-----ILDPLLQ 787
DB 205 RFLSRVHP-----GYRTPWVAILGVSLFSLFALFVLPFRDVAIVANISNFMIF 249

QY 788 MSAAFVALLSLDSKROEASRLDVCCCKVQPELPPGQEGLLGLFFQKAYAPFLLHMITRG 847
DB 250 IIFFMVNLSLIKRLYTDPCRN-----RPFKVP-----LNIGRF--PLLPLF-----G 289

QY 848 VVILLPLALFGVSL-----YSMCHISVGL 871
DB 290 ALSTVFL-FQLSLEVITYGLVFIGIGI 316
```

RESULT 32

```
US-11-096-568A-27723
; Sequence 27723, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27723
; LENGTH: 1080
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)_(1080)
; OTHER INFORMATION: Ceres Seq. ID no. 1825823
US-11-096-568A-27723

Query Match
  1.5%; Score 105; DB 7; Length 1080;
Best Local Similarity 19.7%; Pred. No. 8.6;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

QY 300 ILLVGRVA-PARDKSKMVDPKGTSLSDKLSFSTHTLLGQFFQGWG-----TWVASWP 352
Db 44 VQLIGPVLFLFRFFSQTGNPK--PTEADFRKFSSTILGSVYIVTGLLVANSYMSVG 101
QY 353 LTLVLVSIPVVALAAGLVFTELTDPVLMWSPNSQARSEKAFHDHGFPPFRTNQVIL 412
Db 102 LYLVPVSTSLI-LASQLAFT-----APFSYFLNSQKFTPTIVNSLFL 144
QY 413 TAPN-----RSSRYVDSLLGPNFSGIL-----DLDLLELLEL-QERLRHLQVSP 459
Db 145 TISSALLVNTDSENTAKVRVKYVIGIITIGASAGIGLLSLVQLILRKVLKQKTF 203
QY 460 EAQRNLSLQD-ICYPAPNDNTSLYDCINSLLQYFQNN--RTLLLTANQTLMGQTSOV 516
Db 204 -----TVTDLVAY-----QSLVASCV-VLIGLFASGEWKTLTSEMENYKL-GKVPV 248
QY 517 -----DWKHFLYCANAPLTFKDGITALSCMADYCAPVPFPAIGYKKG-DYSEA 567
Db 249 MTLASIAISQWYVTVGVG--LIFESSVFSNSITA-VGLPIVPVAVIVFHDKNASKI 305
QY 568 EALIM-----TFSLNNY-----PAGDPRLAQAKLWEAFLEEMRAFORRMAG 609
Db 306 FSIILAIWGFISFVYQHYLDEKKLTSHTSVPDGPLLPA---EEG----- 348
QY 610 MFQVTFABRSLE-----DEINRTAEDLPFATSIYIVIFLYISLALGSYSWSRVMVD 663
Db 349 --HTNIHSDQNLEANLIDHEVVTESSSAVP-----QENYKRWLEVSIV 391
QY 664 SKATIGLGGAVALGAVMAAMGFPS-YLGRSSLVILQVVPFLVSVGADNIFIFVLEQV 722
Db 392 VIFVLFCOPLATILGRLYYENGSTYV-----VTLLQLIGPVLV-----LFRP---PS 438
QY 723 RLPRRPGPREVHIGALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTS-----GL 778
Db 439 RIQPKSTDNFSQSPSFTTLASVYLCTGLLVSAVAYLSALA-FTAFPSYFLNSQKFTPL 497
QY 779 AVILDFLLQMSAFVALLSDSKRQ-EASR---LDVCCCKVQELPPPCQGRGLLIGPQ 833
Db 498 IVSSLLLLTVSSALLVNTDSENTSVSRVQVIGICTI-----GASAGIGLLSLIQ 551
QY 834 KAYAPFLHMTIRGVVLLFLALFGVSLYSMCHISVGL-----DOELALPKDS 881
Db 552 MLFRKVFTRKHTSSAVT--DLAIY-QSLVASCVVLIGLFASGEWETLPSEMENYKLGVS 607
QY 882 YLLD-----YPL-----FLNRYFEVGAPVY----- 901
Db 608 YVLTLASAIAISQWYVTVGLVGLIFESSVFSNSITAVGLPIVPEVAIVIFHDMASKIP 667
QY 902 -----FVTLTGYNFSSEAGNAICSSA-----GCNFSFTQKIQ-YATEPPE 942
Db 668 SIILAIWGFISFVYQHYLDEKKLTSHTSVAGDLHLFVEEGHTNIQSDQNLEANLIDHEE 727
QY 943 QSYLAIPASS---WV-----DDFIDWLTTPSSCC--RLYISGPNKDKFCPSVNSLNCILK 991
Db 728 TESFSVPQTKNCKRWLRVSIYALFVFCOPLATVLRGLYYENGSKSYVVTLLQLIGPVP 787
QY 992 NCWSTMGSVR--PSVQEPHKYLPWFLNDRPNKCPKGGLA-----AYSTSVNL-----TS 1040
Db 788 LILFRFSSRIQPKSTDNFSQSPSFTTLASVYLCT--GLLVSAVAYLSAVGLLVPVST 845
QY 1041 DQOVLASR--FMAYHKPLNSQDYTEALRAARELANITADLRKVPGTDPAPEVPPYIT 1098
Db 846 FSLIILASQAFTAFPSYFLNSQKFTPLIVNSLFLTLTVSSALL--VVNTDSE-----NTT 897
```

```
QY 1099 NVFYBOYL-----TI-----LPRGLF-----MLSLCLVPTFANVSC-LL 1130
Db 898 NVSRVQYVIGICTIGASAGIGLVLSLIQLLFRKVKTKHTSASVADLANYSQSLVATCVVL 957
QY 1131 LGL-----DLRSGLNLLSIYVIMLVDTVGMALWDI--SYNAVSLI-----N 1170
Db 958 IGLFASGEWRTLPSEMRNYKLGKVSYILTASAAIF---WQVTVCGVGLIFESSSVFSN 1014
QY 1171 LVSAVGMSV 1179
Db 1015 SITAVGLPI 1023

RESULT 33
US-11-096-568A-27722
; Sequence 27722, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27722
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1097)
; OTHER INFORMATION: Ceres Seq. ID no. 1825822
US-11-096-568A-27722

Query Match
  1.5%; Score 105; DB 7; Length 1097;
Best Local Similarity 19.7%; Pred. No. 8.8;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

QY 300 ILLVGRVA-PARDKSKMVDPKGTSLSDKLSFSTHTLLGQFFQGWG-----TWVASWP 352
Db 61 VQLIGPVLFLFRFFSQTGNPK--PTEADFRKFSSTILGSVYIVTGLLVANSYMSVG 118
QY 353 LTLVLVSIPVVALAAGLVFTELTDPVLMWSPNSQARSEKAFHDHGFPPFRTNQVIL 412
Db 119 LYLVPVSTSLI-LASQLAFT-----APFSYFLNSQKFTPTIVNSLFL 161
QY 413 TAPN-----RSSRYVDSLLGPNFSGIL-----DLDLLELLEL-QERLRHLQVSP 459
Db 162 TISSALLVNTDSENTAKVRVKYVIGIITIGASAGIGLLSLVQLILRKVLKQKTF 220
QY 460 EAQRNLSLQD-ICYPAPNDNTSLYDCINSLLQYFQNN--RTLLLTANQTLMGQTSOV 516
Db 221 -----TVTDLVAY-----QSLVASCV-VLIGLFASGEWKTLTSEMENYKL-GKVPV 265
QY 517 -----DWKHFLYCANAPLTFKDGITALSCMADYCAPVPFPAIGYKKG-DYSEA 567
Db 266 MTLASIAISQWYVTVGVG--LIFESSVFSNSITA-VGLPIVPVAVIVFHDKNASKI 322
QY 568 EALIM-----TFSLNNY-----PAGDPRLAQAKLWEAFLEEMRAFORRMAG 609
Db 323 FSIILAIWGFISFVYQHYLDEKKLTSHTSVPDGPLLPA---EEG----- 365
QY 610 MFQVTFABRSLE-----DEINRTAEDLPFATSIYIVIFLYISLALGSYSWSRVMVD 663
Db 366 --HTNIHSDQNLEANLIDHEVVTESSSAVP-----QENYKRWLEVSIV 408
QY 664 SKATIGLGGAVALGAVMAAMGFPS-YLGRSSLVILQVVPFLVSVGADNIFIFVLEQV 722
Db 409 VIFVLFCOPLATILGRLYYENGSTYV-----VTLLQLIGPVLV-----LFRP---PS 455
QY 723 RLPRRPGPREVHIGALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTS-----GL 778
```

Db 456 RIQPKSTDTNFSQSPSTTLASVYLCTGLLVSAAYALSALA-FTAFSSYFLNSQKFTPL 514
Qy 779 AVILDFLOMAFAVALLSLDSKQ--EASR-----LDVCCVKPQELPPPGQEGLLGFPQ 833
Db 515 IVSSLLLTVTSALLVNDSENSTNVRVQVIGICTI-----GASAGIGLLLSLIQ 568
Qy 834 KAYAPFLHWTIRGTVLLFLALFGVSLYSMCHISVGL-----DQELALPKDS 881
Db 569 MLFRKVTHTSSAVT---DLAIY-QSLVASCVVLIGLFGSEWETLPSEMRNYKLGKVS 624
Qy 882 YLLD-----YFL-----FLNRYFEVGAPEVY----- 901
Db 625 YVLTLASAAISQVYTLGLVGLIFESSVFSNSITAVGLPIVPAVAIVFHRMDASKIP 684
Qy 902 -----FVTLGYNFSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db 685 SIILAIICGFLSFVYQHYLDEKCLNTSHTSAVGDHLHPVEEGHTNQSDQNLLENDHEE 744
Qy 943 QSYLAIPASS-----WV-----DDFIDWLPSSCC--RLYISGPNKDKFCPSTVNSLNCIK 991
Db 745 TESFSVPQTKNCKRMLRVSIYAIFFVFCOPLATVGLRLYYENGSGKSTYVVTLLQIGFPV 804
Qy 992 NCWSTMTGMSVR--PSVEQPHKLPWFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db 805 LILFRFSSRIQPKSTDTNFSQSPSTTLASVYLCT--GLLVSAAYALSAVGLLXLPVST 862
Qy 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPEVFPYIT 1098
Db 863 FSLILASQALFTAFSSYFLNSQKFTPLVNSILFLTVSSALL--VVNTDSE-----NTT 914
Qy 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db 915 NVSRVQVIGICTIGASAGIGLVLSLIQLLFRKVTHTSSAVLDLANYSQSLVATCVVL 974
Qy 1131 LGL-----DLRSGLLNLLSIWMLVDTVGFMAWLDI--SYNAVSLI-----N 1170
Db 975 IGLFASGEWRTLPEMRNYKLGKVSILTASAAIF---WQVYTVGCVGLIFESSSVFSN 1031
Qy 1171 LVSAGMSV 1179
Db 1032 SITAVGLPI 1040
RESULT 34
US-11-096-568A-27721
; Sequence 27721, Application US/110956568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27721
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1128)
; OTHER INFORMATION: Ceres Seq. ID no. 1825821
US-11-096-568A-27721

Query Match 1.5%; Score 105; DB 7; Length 1128;
Best Local Similarity 19.7%; Pred. No. 9.1;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

Qy 300 ILLVGRVA-PARDKSKVDPKGTSLSDKLSFTSTHTLLGQFGWG-----TWVASWP 352
Db 92 VQLIGFPVLFRFPSTKPK--PTBADFRKSSFTILGSVIYIVTGLLVSAANSYMSVG 149

Qy 353 LTIIVLSVIPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHFQPPFRTNQVIL 412
Db 150 LLYLPVSTFSLI-LASQALFT-----AFFSYFLNSQKFTPIVNSLELL 192
Qy 413 TAPN-----RSSRYDLSLLGPKNFSGIL-----DLDLLELLEL--QEBRLHLQVWSP 459
Db 193 TISSALLVNTDSENTAKSVKVIIGLICIGASAGIGLLLSLVQLLRLKVLKQTF-- 251
Qy 460 EAQRNLSQD--ICVAPLNPNTSLYDCINSLLOYFQNN--RTLLLTANQTLMGQTSQV 516
Db 252 -----TVTDLVAY-----QSLVASCV-VLIGLFASGEWKTLTSEMENYKL-GKVPYV 296
Qy 517 -----DWKDHFLYCANAPLTFKDGCTALALSCHADYGAPVFPFLAIGGYKKG--DYSEA 567
Db 297 MTLASIAISQVYTVGVG--LIFESSVFSNSITA-VGLPIVPAVAIVFHDKNASKI 353
Qy 568 EALIM-----TFSLNYY-----PAGDPRLAQAOLWEBAFLERAFORRMAG 609
Db 354 FSIILAIWGFISFVYQHYLDEKCLNTSHTSPVGDPHLLPA---EEG----- 396
Qy 610 MFQVTFTRASLE-----DEINRTAEDLPFATSYIVIFLYISLALGSYSSWSRVMD 663
Db 397 --HTMIHSDQNLLEANLIDHEVVTSESSAVP-----QTENYKRWLRVSIY 439
Qy 664 SKATILGLGVAVVLGAVMAANGFES-YLGISSSLVLOVVPFLVLSVGDNIFIFVLEYQ 722
Db 440 VIFVFCOPLATILGRLYYENGSTYV-----VTLQLIGFPVLV---LFRP---PS 486
Qy 723 RLPRRPGPREVHIGRALGRVAPSMLLCSLSEACIFFLGALTPMPAVRTEALTS-----GL 778
Db 487 RIQPKSTDTNFSQSPSTTLASVYLCTGLLVSAAYALSALA-FTAFSSYFLNSQKFTPL 545
Qy 779 AVILDFLOMAFAVALLSLDSKQ--EASR-----LDVCCVKPQELPPPGQEGLLGFPQ 833
Db 546 IVSSLLLTVTSALLVNDSENSTNVRVQVIGICTI-----GASAGIGLLLSLIQ 599
Qy 834 KAYAPFLHWTIRGTVLLFLALFGVSLYSMCHISVGL-----DQELALPKDS 881
Db 600 MLFRKVTHTSSAVT---DLAIY-QSLVASCVVLIGLFGSEWETLPSEMRNYKLGKVS 655
Qy 882 YLLD-----YFL-----FLNRYFEVGAPEVY----- 901
Db 656 YVLTLASAAISQVYTLGLVGLIFESSVFSNSITAVGLPIVPAVAIVFHRMDASKIP 715
Qy 902 -----FVTLGYNFSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db 716 SIILAIICGFLSFVYQHYLDEKCLNTSHTSAVGDHLHPVEEGHTNQSDQNLLENDHEE 775
Qy 943 QSYLAIPASS-----WV-----DDFIDWLPSSCC--RLYISGPNKDKFCPSTVNSLNCIK 991
Db 776 TESFSVPQTKNCKRMLRVSIYAIFFVFCOPLATVGLRLYYENGSGKSTYVVTLLQIGFPV 835
Qy 992 NCWSTMTGMSVR--PSVEQPHKLPWFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db 836 LILFRFSSRIQPKSTDTNFSQSPSTTLASVYLCT--GLLVSAAYALSAVGLLXLPVST 893
Qy 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPEVFPYIT 1098
Db 894 FSLILASQALFTAFSSYFLNSQKFTPLVNSILFLTVSSALL--VVNTDSE-----NTT 945
Qy 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db 946 NVSRVQVIGICTIGASAGIGLVLSLIQLLFRKVTHTSSAVLDLANYSQSLVATCVVL 1005
Qy 1131 LGL-----DLRSGLLNLLSIWMLVDTVGFMAWLDI--SYNAVSLI-----N 1170
Db 1006 IGLFASGEWRTLPEMRNYKLGKVSILTASAAIF---WQVYTVGCVGLIFESSSVFSN 1062
Qy 1171 LVSAGMSV 1179
Db 1063 SITAVGLPI 1071

RESULT 35

US-11-087-099-9820

Sequence 9820, Application US/11087099

Publication No. US20060041961A1

GENERAL INFORMATION:

APPLICANT: Abad, Mark S. et al.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53450)B EP

CURRENT APPLICATION NUMBER: US/11/087,099

CURRENT FILING DATE: 2005-03-22

NUMBER OF SEQ ID NOS: 12464

SEQ ID NO 9820

LENGTH: 315

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

NAME/KEY: unsure

LOCATION: (1)-(315)

OTHER INFORMATION: unsure at all Xaa locations

US-11-087-099-9820

Query Match 1.5%; Score 104.5; DB 7; Length 315;

Best Local Similarity 22.7%; Pred. No. 1.6;

Matches 52; Conservative 40; Mismatches 82; Indels 55; Gaps 11;

QY 1088 PAPEVFPYTTINVPYEOYLITLPEGLFML-----SLCLVPTFAVSCLL--LGLDL 1135

Db 54 PSYLPPELVKFPFFAN--LSIVGMNVSIMMNSVGFYQIAKLCIIP---VLCILEILFDKV 109

QY 1136 RSGLLNLLSIVMLIVDTVGFWALWDIYNVAVSLINLVSAV---GMSVEFVSHITRSPAIS 1192

Db 110 RYSRNTKLSIVLVVG-VAVCTVTDVSNVSKGLLAAVIAVMTALQQHYVHHLQKYSLG 168

QY 1193 -----TKPT-----WLERAKEATISMGSAVFAGVAM-----TNLPGI 1224

Db 169 SPNLLGHTAQAASLLILGDFVDFWUTNRVDVTFNTTIVTFIVLSCTIAVGTNLNQF 228

QY 1225 LVLGAKAQLIQFFPRINLLITLGLL-----HGL---VFLPVLISYVG 1266

Db 229 ICIGRPTAVSFQVLGHMKTIILVLTGLPFGKEGILNFHVLGMMLAVIG 277

RESULT 36

US-11-087-099-4362

Sequence 4362, Application US/11087099

Publication No. US20060041961A1

GENERAL INFORMATION:

APPLICANT: Abad, Mark S. et al.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53450)B EP

CURRENT APPLICATION NUMBER: US/11/087,099

CURRENT FILING DATE: 2005-03-22

NUMBER OF SEQ ID NOS: 12464

SEQ ID NO 4362

LENGTH: 471

TYPE: PRT

ORGANISM: Bacillus cereus ATCC 14579

US-11-087-099-4362

```

Db      188 E-GTISLSAILQVKNMLV-----TLWVFIG-----VEGAVVLSGRAKNSRDVGKAT 234
QY      781 ILDFLLQMSAFVALLSLDSKQEQASRLDVCCCVKQPELPPQGE-----825
Db      235 VLGILVMSIYILISVLISMGAMTRGELSV-----LETFSMGHVLEHVVGPMGAVAIN 286
QY      826 -----GLLGLGPF-----QKAYAPFLHWTIRGVVLLLF- 853
Db      287 IGLVASLVGTLIGWFLVSBISHVAGXGVPPKVTNKQTPTMALWISNGVAQIIFI 346
QY      854 LALFGVSLYSMCHISVGLDQELALPKDSYLLDYFL 888
Db      347 IVLFSBSTYQIMYP-----IASTSILLPYLL 372

RESULT 37
US-11-096-568A-27509
; Sequence 27509, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polynucleotide Sequences
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27509
; LENGTH: 757
; TYPE: RPT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(757)
; OTHER INFORMATION: Ceres Seq. ID no. 1818832
US-11-096-568A-27509

```

Query Match	1.5%; Score 104; DB 7; Length 757;
Best Local Similarity	20.6%; Pred. No. 6.2;
Matches	145; Conservative 91; Mismatches 209; Indels 260; Gaps 41;
Qy	670 LGGVAVVLGAVMAAG-----PFSYLGIRS-----SLVI---LQVP-----FL 705
Db	20 LGIAFOTLGVVYGDMTSPLYVSDVFSEKPIRSEVDVLGALSIVYITIAVPLAKYFV 79
Qy	706 VLSV-----GADNIFVLEVYQBLPRPG-EPREVHIG-----RALG-----741
Db	80 VLKANDNGEGNASAPMCLVKYAKVKNLPNOQDAEQISSFLKLTPTPELGRALIKRALE 139
Qy	742 --RVAPSMLLCSIEAICFFLIGALTMPAVRFTALTSGL-AVILDF---LLOMSAFVALL 795
Db	140 TKGYLTKTLLLLVLMTGSMIIGDILTPAMSVMSAGLQGEVKGFGCTNALVMSSIVILV 199
Qy	796 SLDSKQOBSRLDVCCVKPQELPPGQGBGLLLGPFQKAYAPLLHWTGRGVVLLFLA 855
Db	200 ALFSIQRF-----GTGK---VGFL---FAPVLALW-----FFS 226
Qy	856 LFGVSLSYMSCHISVGLDQELALPKDSVLLDYFLFLNR-----YFEGVAPYFVT-----904
Db	227 LGAIGNILNKYDFTVIRAL---NPPIYV---LFFNKNKQAWSSALGGCVLCITGABAMF 280
Qy	905 -TGYNFSSEAGMNAICSSAGCNFNFSFTOKIQTATBPQEOS-----YLAIPAS-SWVDDFI 958
Db	281 ADLGHSVRSIQMAFTCVVFPCLLLAYMGQAAVLTKHPEASARIFVDSVPKSLFWPVPVI 340
Qy	959 DMLTPSSCCRLYISGPNKDKFCBSTVNSLNCNCSITMGSVRPSVEQPHK-----1010
Db	341 ATLAAMIASQAMISA-----TFSCVQAMAL---GCFPLKLIHTSKRIGQI 385
Qy	1011 YLP---WFLNDRPNIKCPKGLAAYASTVNLSTSDGOVLASRFMAYHKPLKNSODYTEALR 1067
Db	386 YTPVINWFL---MINC-----ILV-----VSIFR 406

Query Match	1.5%	Score 104.5;	DB 7;	Length 471;
Best Local Similarity	20.0%;	Pred. No. 2.9;		
Matches	67;	Conservative 42;	Mismatches 101;	Indels 125; Gaps 15;
QY	643	IPLYISLALGSY-----SSWSRVVDSKATLGL-----GG--VAVVLGA- 679		
Db	74	IYSYARAGEYIGFNSANCYWLAILGNVATTMLLFSTLYGFFPIKGGNVASIVGAS 133		
QY	680	-VMAAGFFSYLGIIRSLV-----ILQVVPFLVLSVGADNIIFVLEYORLPRRP 729		
Db	134	LLLTWTLHFLILFGIREASIMNVATTIGKLVPIVLFIWMVATFRWTFTTHDFW-----G 187		
QY	730	EPREVHIGRALGRVAPSMLLCSLSEACPFGLGALTMPAPVTRPALTS-----GLAV 780		

[illegible]

RESULT 40

US-11-087-099-1098
; Sequence 1098, Application US/11087099
; Publication No. US20060041961A1

```

; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464

```

US-11-087-099-1098

	Query Match	1.5%; Score 103.5; DB 7; Length 251;
	Best Local Similarity	24.3%; pred. No. 1.4;
	Matches	53; Conservative 35; Mismatches 79; Indels 51; Gaps 10;
QY	1097	ITNVFYEQYLTIILPEGLFMLE-----SLCLVPTFAVSCLLLGDLRSGLNILLSIV 1146 : : : : : 1 ITLVFFAN-LSVGMMNSLMMWSVGFYQIAKLSIPLLCIMEVLFE-NFRYSRDTKLISIV 58
Dd	1147	MILDVTGVFWALWDISYNASVINLNLAV---GMSVEFVSHITSFAIS-----TKPT 1196 : : : : : 59 VVLVG-VGVCTSDSVSNAOQLVAATVCGTALQQHVNYVLKYKSINSLKLGHTAPA 117
QY	1197	-----WLRKEATTISMGSAPFAGVAM-----TNLPGLIVLGLAKAQLI 1235
Dd	118	QAASLLILGPFDVPFLNRNIDTFHYSTWTFTFVLSVCISVSGNTSQFCICIGRTAVTF 177 : : : : : : : : : :

RESULT 41

US-11-087-099-8831
; Sequence 8831, Application US/11087099
; Publication No. US20060041961A1

```

; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464

```

; ORGANISM: Oryza sativa
US-11-087-099-8831

Query March	1.58; Score 103.5; DB 7; Length 331;
Best Local Similarity	22.7%; Pred. No. 2.1;
Matches 52; Conservative 40; Mismatches 82; Indels 55; Gaps 11;	
Qy	1088 PAFEVPTTITNVFEQYLITLPEGLFML-----SILCVPTFAVSCLL--LGLDL 1135
Db	70 PSYLPPLVKVFPPAN-LSIVGMNVSIMWNSVCFYQIAKLCIIP---VLCILEILDPKV 125
Qy	1136 RSGLLNLLSIWMLIVDTVGPMAWDISYNAVSLINLVSAV---GMSVEFPVSHITRSPAIS 1192
Db	126 RYSRNTKLSIVLVAG-VAVCTVTDVSNVSRGILLAAVIAVWSTALOQYVHHLQKCYSLG 184
Qy	1193 -----TKPT-----WLERAKEATISGSAVPAQVAM-----TNLPGI 1224
Db	185 SPNLLGHTAPQAASLLILGPFVDVFWLTNRVRDVTFNVTTFPIVLSCTIANGVTNLSQF 244
Qy	1225 LVIGLAKAQIUIQFFFFRNLNLTITLGLL-----HGL-----VFLPWLVSYG 1266
Db	245 ICIGRFTAVSFOVLGHMKTTILVLTGLFPFGKEGLNFHVHVLGMLVAVIG 293

RESULT 42

US-10-506-454-134
; Sequence 134, Application US/10506454
; Publication No. US20060068386A1

```

: GENERAL INFORMATION:
:
: APPLICANT: Slesarev, Alexi I
: APPLICANT: Mezhevaya, Katja V
: APPLICANT: Polushin, Nikolai N
: APPLICANT: Shcherbinina, Olga V
: APPLICANT: Shakhova, Vera V
: APPLICANT: Malykh, Andrei G
: APPLICANT: Kozayavkin, Sergei A
:
: TITLE OF INVENTION: The Complete Genome and Proteome of the
: Fungus Methanopyrus kandleri AV19
:
: TITLE OF INVENTION: Methods and Methods of Use Thereof

```

; ORGANISM: Met
US-10-506-454-134

Query Match	1.5%;	Score 103.5;	DB 6;	Length 2042;	
Best Local Similarity	18.5%;	Pred. No. 28;			
Matches 160;	Conservative 108;	Mismatches 324;	Indels 275;	Gaps 39;	
QY	32	YCAFYDBC-----	GKNPELGGSLMTLSNVSCLSNTPARKITGDHLLILQLKICPRLY	82	
Db	1239	YIVLTGCTEVRVRIDBGRVPEGEYTVIGTV-----	LVGDEPVI-----	1280	
QY	83	TGNTQAQCSAKOLVSLEASLSITKALLTRCPACSNFVNLHCHNTCSPNQSLFINVTRV	142		
Db	1281	-----SASLVLRVPRRLTIVTGTSCFRDVF-----	AHNT-----	L 1313	
QY	143	AQLGAGOLPAWVAAYEAPYQHSFAEQS-----	YDSCSRVRPAAATLAVGTCMGVYGSAALC	197	
Db	1314	TELFELGRVAAILAAEYFLARLLSHSLIYSDVDKVLIR-----	GIDEPYVY	1360	
QY	198	NAQRWLN-----	FQDGTNGNGIAPL-DITFH-----	LLEPGQAVSGGIQPLNEGVA	241
Db	1361	WVRWMLRETVERLSQCEWKN-----	PIHETVLHVYRAALWRAALTTPDEBLA----	RELEGR	1414

```
QY 242 RCNESQDDVATCSCODCAASCPAARQALDSTFYLQMPGSLVLIILCSVFAVVTL 301
Db 1415 RLYDLVGE-----FCAEHGYDFESVNTAA-----SLAVLPFAENYFPLDWA 1457
QY 302 LVGFVAPARDKSMVDPKKTSLSKLSFSTHLLGQFFQGWGTWVASWPLTIL-VLSV 360
Db 1458 AVYLLALLCRESNITSAADAGSGSPN-----GGIITGWIEQVREVDEDLPLAA 1508
QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHQHFGPPPRTRNOVILTAPNRSSY 420
Db 1509 LNVFQCAPG-DYTNVIGKTIE-----SGEFLLEDRLNALALSWISGL 1548
QY 421 RYDLSLLGPNFSGILDLDLLELELOERLHLQWSPEAQRNISLQDICVAPLPDNT 480
Db 1549 SY---VYGPKNWAGSFPLLALN-LAAPDRTLHTWVSDEKA-----T 1587
QY 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGTAAL 540
Db 1588 PFYDDCIYA-----FEGGLRLAVSAVNGLLPQETMID-----ALVL 1624
QY 541 SCWADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPRLAQKLMEEAFLEEM 600
Db 1625 NLRN-----AALGSYVGDYSQM-ARRLAEBITQLMAAANPNLAP---WLRDLASOL 1671
QY 601 RAPQRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYYIVIELYISLALGSYSWSRV 660
Db 1672 -----SAGSYLAALIAS-NMTLSFTVTRTALATGNETAARFYSLLMPDPTAWYDL 1724
QY 661 MVDSKATLGGVAVVLGAV-MAAMGFFSYLGRSLVILQVVPFLVLSVGADNIFIVL 719
Db 1725 M-----RTVFNVPVYVGLRFHGYSG---AVELLKRLGYLIR--GWSTLALGLL 1767
QY 720 BYQRLPRRPG-----EPREVIHGRALGRVAPSMLLCSLSEACFFLGLALTPMAVTRFALT 775
Db 1768 GWEGIPRRTASTLVNRE-----WLARYSGEGL-----FSLA 1799
QY 776 SGLAVL-----DFLLQMSAFVALLSLDSKQEASRLDVCCC-----VKPQE 817
Db 1800 VSLLYAYDRAQHKLSEELVSEFVALVR-DVIVPELLRGLVCCCGVCGNVPVQORL 1858
QY 818 LPPPGQGGELLGFFOKAVAPFLHVI 844
Db 1859 LEALSQYGLVFN-LREANAVFANVM 1884

RESULT 43
US-11-126-313-34
; Sequence 34, Application US/11126313
; Publication No. US20050288489A1
; GENERAL INFORMATION:
; APPLICANT: Hirsch, Joel
; TITLE OF INVENTION: VOLTAGE-DEPENDENT CALCIUM CHANNEL BETA SUBUNIT FUNCTIONAL CORE
; FILE REFERENCE: P-6758-US
; CURRENT APPLICATION NUMBER: US/11/126,313
; CURRENT FILING DATE: 2005-05-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34
; LENGTH: 2312
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-126-313-34
Query Match 1.5%; Score 103.5; DB 7; Length 2312;
Best Local Similarity 18.2%; Pred. No. 33;
Matches 226; Conservative 161; Mismatches 436; Indels 421; Gaps 59;

QY 125 CHNTCSGNQSLFNITVRVAOLGAGQLPAAVVAEAFVQHSFAE-----QSDSC 172
Db 887 CHGNCOPTQ-----BAGGGE--AVVTFEDRAHRQSORRRRRVRTEGKSSA 935
QY 173 SRVRVPA--AATLAVGT-----MCGVYGSALCNAQ-----RWLNFQDGTGNGLA 214
```

```
Db 936 SRSASQERSLDEAMPTEGEKDHRLGNHGAKEPTIQEERAQDLRRTNSLMVSRGSLA 995
QY 215 -----PLDITFHLLPQAV-----GSGIQLNEGVAR-----CNESQGD--- 249
Db 996 GGLDEADPLVLPHELEVEGVHVVLTQEOPGSGEQALLGNVQDMDGRVISOSEDLSCI 1055
QY 250 ----DVATCSQDCAASCPAARQALDSTFYLQMPGSLVLIILCSVFAVVTL--- 301
Db 1056 TANTDKATTESTSVTVAIPDV--DPLVDST-----VVHISNKTD 1092
QY 302 ---LVGFRVAPARDKSMVDPKK-----GTSL---SDKLSFST-----HTLLG-Q 339
Db 1093 GEASPLKEAREDEEBEVEKKQKKEKRETKAMVPHSSMFISTTNPTRRACHYIVNLR 1152
QY 340 FQOQGTWVASWPLTILVLSVIPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHQ 399
Db 1153 YFE-----MCILLVIAASSIALAA-----BDPV-----LTNSRNVKLYRFDY 1190
QY 400 HFGPPFRTRNOVI-----LTAPNRSSRYDSILLGPKNFGSLDLDLLELL----- 445
Db 1191 VFTGVTFTEMVIMKIDQGLILQDGSYFR-----DLWNILDFVVVVVGVAVPALANA 1241
QY 446 -----ELQERLHLQWSPEAQRNISLQDICVAPLPDNTSLYDCCINSLLQYFQN 496
Db 1242 LGTNKGRDIKTIKSLRVLRLRP-----LKTI---KRLPKLKAFCVVTSLKNVF-- 1289
QY 497 NRTLLLTANQTLMGQTSQVDMKDHFLYCANAPL-TFKDGTALALSCWADYGAPVFPPLA 555
Db 1290 -NILIVYKLFMFIFAVIAVQVLFKGFYCTDSSKDKTEK-----C 1328
QY 556 IGGY-----KGKDYSEAE-----ALIMTFSLNYPAGDPRLAQKLMEEAFL 597
Db 1329 IGVYDHEKKNKEVKGREKREHFEHYDNIIWALLTLFTVST-GEQVQVLQ----- 1378
QY 598 EEMRAPORMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYYIVIF-----LYISLAL 651
Db 1379 -----HSDVTEEDRGPSSNRM---EMSIFYVYVYVVFVFFVFIFVALLII 1422
QY 652 GSYSSWSRVWDSKATLGLGVAVVLGAVVAVMAAMGFFSYLGRS-----SILVILQVVP 703
Db 1423 ITFQSGQKQMBE-ECSLEKNERACIDFAI-SAKPTRYNPQNRHTFOYRWVHVVSPTSFE 1480
QY 704 FLVLSVGADNIFIVLEYQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFFLGLAL 763
Db 1481 YTIMAMIALNTVLMKKYISAP-----CTYELALKYLNIAF 1516
QY 764 TPMPAVTRFALTSGLAIV-----ILDPLLQMSAFVALLSLDSKQEASR--- 806
Db 1517 T-----MVPSLECVLKVIAFGFLNVPFRDTWNIFDFITVIGSITEIILTDSKLVNTSGFNM 1571
QY 807 --LDVCCCVKPOELPPPGQEGELLGFFOKAVAPFLHVIITRGVLLLFL-ALFGVSLYS 863
Db 1572 SFLKLFRAARLKLRLQGGVTIIRLLMTFVQSPK--ALPVCLLIAMLFFIYALIGQVFG 1629
QY 864 MCHISVGLDQELALPKDSYLLDYF-----LFLNRYFEVGAFFVFTTL 906
Db 1630 ----NIKLEESHINHNFRSFFGSLMLLFRSATGEAMQEIIMSLGKGCCEPDTTAPS 1685
QY 907 GYNFSSEAGMNA-----ICSSAGCNPF--SFTQKIQYATEFPEQSYLAIPASSWV 954
Db 1686 GQNERECGTDLAYVYVVSFIFPFCSPFLMLNLFVAVIMDNFEXLTR--DSSIL---GPHHL 1740
QY 955 DDFID-W--LTPSSCCRLY-----ISGP-NKDKFCPSVNSLNCNKMCSITMGS 1000
Db 1741 DEFVRVWBYDRAACRIIHYTEMYEMLTMSPLPLGLGKRCPSKV-----AYKRLVLMMPV 1796
QY 1001 VRPSVEQFHKLYLPWFVLDNRPNIKCPKGG-----LAAY 1032
Db 1797 AEDMTVHFTSTLMALIRTALDIKAKGADRQQLDSELQKETLAIWPHLSQKMLDLLVPM 1856
QY 1033 STSVNLTSDQGVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPGTDDPA--- 1089
```



```
Db 1857 PRASDLTVGKIYAMIMDYK-----QSKVKQROQLBEQKNAPMFORWEPSSLQEI 1911
QY 1090 --FEVFPYITNVF-----YEQYLITLPEGLFMLSCLVPT 1123
Db 1912 ANAKALPYLQODPVSGLSGRSGYPSPLSPQDIFOLA-CMDPT 1954

RESULT 44
US-11-087-099-3980
; Sequence 3980, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3980
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Pseudomonas putida KT2440
US-11-087-099-3980

Query Match 1.5%; Score 103; DB 7; Length 463;
Best Local Similarity 20.0%; Pred. No. 3.7;
Matches 130; Conservative 66; Mismatches 183; Indels 270; Gaps 32;

QY 203 LNFQGTGN---GLPLDITF-----HLLBPQAVGSGIQPLNEGVARCNEQDDVATCS 255
Db 1 MNTVGSQDGLAQGFKPRHVTMLSGIAGLGVGSG----- 37
QY 256 QDCAAACPAIARPOALDSTFYLGQMPGSLVLIILCSVFVAVVTILLGVFRVAPARDKSK 315
Db 38 -----HATAAGPATIISYFVAGTL---VLVWRMLGEMAV----- 70
QY 316 MVDPKGTSLSKLSSTHTLQPFQGWGTWVASWPLFILVLSVIPVVALAAGLVFTL 375
Db 71 -AHPDTG-----SFTSYA--DQAIGRWAGYTIGLWFWLVIPIEALAAAGHVLNAW 120
QY 376 TTDPVELWAPNSQARSEKAFDQHEGFPFRITNOVILTAPNRSSYYDSLLGPKNFSGI 435
Db 121 FPQ-VDSW-----IFALASVILLA----- 138
QY 436 LDLDLLELLEQLERLRLHQLVQWSPQAQRNLSQDICYAPLNDNTSLYDCCINSLLQYFQ 495
Db 139 -----GTNLP-----SVAKYGE 150
QY 496 NNRTLLLLTANOTL-----NGQTSQVDWKDHFLYCANAPLTFKDGTTALALSCMADYG--AP 549
Db 151 FEFWFALIKVTAITGLFGLFAALMDW-----LPNREVSGLS-GLMAEYGGFAP 198
QY 550 -----VPFPL-----AIGYKGYSE--AEAL-----IMTF-----SL 576
Db 199 KGMSAVVGAFITVMFSGITGEATVIAAESDPSRNIKAATRSVIRISTFYILSFVII 258
QY 577 NNPVAGDPRIQAQKLWEEAFLEBMR---AFQRMAGMFQVTFTAERSLEDEINRTAED 632
Db 259 SVVPWMDPQLAVVGSYQRA-LEIMNTPNAAFWMDLVVLVAVTSCMNSSI----- 306
QY 633 LPFATSYIVIFLYISLAL-----GSYSSSRVMDSKATLGLGGVAVVLGAVMAAGGFF 687
Db 307 -----YIASRMWFLSKRGDAPAFLNKTSKGVFPRAAVFG-----STLIGRAIAVLNYP 355
QY 688 SYLGIRSSILVQVPLVLSVGDNIPTF-VLEYQRLPRRPGEPREVIHIGRALGRVAPS 746
Db 356 APRG-----VFEEFLASSGAIALLVMVIAISQLRMR-----RRERENTE 396
QY 747 M-----LLCSLSAICFEL-GALTPW---PAVRT-FALTSGLAVIDEL 785
Db 397 LKFRMWLFPLYTAVIIFAGALANVMYTPHRAEVSSTGLAIVISFL 445
```

RESULT 45

```
US-11-087-099-2723
; Sequence 2723, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2723
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Bacillus anthracis str. Ames
US-11-087-099-2723

Query Match 1.5%; Score 102.5; DB 7; Length 475;
Best Local Similarity 19.7%; Pred. No. 4.2;
Matches 66; Conservative 43; Mismatches 101; Indels 125; Gaps 15;

QY 643 IFLYISLALGSY-----SSWSRVMDSKATLGL-----GG--VAVVLGA- 679
Db 78 IYSYARAGEEYIGFNSAWGYWLAGILGNVATIMLLFSTLGYFPFPGKNSVASIVGAS 137
QY 680 -VMAAMGFPSYLGIRSSLV-----ILQVVPFLVLSVGDNIPTFVLEYQRLPRRPG 729
Db 138 LLLNTLHFLILFGIREASIMNVIAITIGLVPVILFIVVMVTAFRWDTFTQDFW-----G 191
QY 730 EPREVIHIGRALGRVAPSMLLCSLSAICFPLGALTMPAVRTFALTS-----GLAV 780
Db 192 E-GTISVSISILQVKNMLV-----TLWVFIG-----VEGAVVLSGRAKNSRDVGKAT 238
QY 781 ILDFELLQMSAFVALLSLDSKQREASRLDYCCVCKPOELPPGQGE----- 825
Db 239 VLGLILVMSIYIILISVLMSGANTRGELSV-----LETSPSHVLEHVVTGAVAIN 290
QY 826 -----GLLIGFF-----OKAVAPFLILLHITRQVWLLLF- 853
Db 291 IGLVASLVGTLLGWFLLVSEISHVAGKGVFPKVFTKTKQTPHMAIWSNGVAQVIFI 350
QY 854 LALPGVLSYMSCHISVGLDQELALPKDSYLLDYFL 888
Db 351 IVLPSESTYQIMYF-----IASTSILVPYLL 376

RESULT 46
US-11-087-099-5717
; Sequence 5717, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5717
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Halobacterium sp. NRC-1
US-11-087-099-5717

Query Match 1.5%; Score 102; DB 7; Length 324;
Best Local Similarity 22.5%; Pred. No. 2.7;
Matches 81; Conservative 41; Mismatches 116; Indels 122; Gaps 16;

QY 526 ANAPLTFKDGTTALALSCMADYGVAPVFPFLAIGYKGYSEABALIMTFSLNYPAGDPR 585
Db 7 AAGAAGPAGQAVVAASAGAPVGPQPVFLAVSG-----AL----- 41
QY 586 LAQAKLWEEAFLEEMRAFORMMAGMFQVTFTAERSLEDEINRTAEDLPFATSYIVIFL 645
```

```
Db 42 -----LAGIMRTLF-----ALRDAPVLGS-VVVVL 66
Qy 646 YISLALG-----SYSSNSRVMDSKATLGLGGVAVVLGAV-MAAMGFSSYLGRSSLV- 697
Db 67 WVTVALAPLAGVPTVAAMERVLVALAVTAFGWSRALGATSIAGMLTGVFLGLLAVVLG 126
Qy 698 -----ILQVPEPLVSVGADNIFIFVLEQO-RLPRRPGEPREHIGRA--LGRVAPSMLL 749
Db 127 GYGWFFVLVAFVAVGLATK-----FKYELKADRGVAPNDGARGTANVLONSAAALIA 180
Qy 750 CSLSEAIACFFLALPMPAVRFTALTSLGLAVILDFLLQMSAFVALLSDSKRQESRL-D 808
Db 181 LVLYAA-----HAHP-----FLFAYAGSVATALADTLLSSEVGGLFD 222
Qy 809 VCCCKVQQLPPPG-----QGEGLLGGFFOKAVAPFLHWHITRGVVLLFLALFGVSL 861
Db 223 TPRLVTTFRVDPGTGDVATWQGE-----LAGVVGATIIAVLSVAVFGVTL 268

RESULT 47
US-11-087-099-9244
; Sequence 9244, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9244
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-11-087-099-9244

Query Match 1.5%; Score 101; DB 7; Length 647;
Best Local Similarity 19.2%; Pred. No. 8.5;
Matches 137; Conservative 91; Mismatches 231; Indels 254; Gaps 36;

Qy 645 LYISLALGSSYSSNSRVMDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSLVILQVVPF 704
Db 30 LYLISLGLIIGSGW--LFASSLSTAAYAGGAAILSWIAG-----ILVMFVGL 73
Qy 705 LVLSVGADNIFIFVLEQRLPRRPGEPREHVI--GRALGRVAPSMLLCSLSEALCFGLA 762
Db 74 AYAEIGA-----AIPKSGGITRYPHYTHGGLVYII-----TWAYFLSA 112
Qy 763 LTPMPAVRFTALTSLGLAVILDFLLQMSAFVALLSDSKRQESRLDVCCKVQQLPPPG 822
Db 113 AS-VPAIEAANAAYEYGSYPQLITSGTF-----DGTVTI-----LTPLG 152
Qy 823 QG-EGLLLGFQKAYAPFLHW-----ITRG-----VLLFLALFGVLSYMSCH 866
Db 153 IGLAGLLLIFF-----FFLNYFGVNLGKVTHGAGWKLLVPTITFLALLDLHSA-- 204
Qy 867 ISVGLDQELALPKDSYLLDYFLFNRYFEVG-----APYFVFTLLGYNFSSEAGWN 917
Db 205 -----NFTLGGGFFPSAQYVYKGGSGYIGFSAVLPAIPSTGVIFA----- 244
Qy 918 AICSSAGCNFFSFTKIQAYTEFPEQSYLAIPASSWVDDFDLWLTSSCCRLYISGNKD 977
Db 245 -----YLFQRAVEYGE-----GKNPS 262
Qy 978 KFCEP-STVNSLNCILKNCISITMGSVRSVQFHKYLFWFLNDRPNKCP--KGGLAYST 1034
Db 263 KDIPFVAVLGSLLIAIALYTLTQVFSIGGIDWSKLYL---VNKTIGILIPVVPNGWSALST 319
Qy 1035 SVNLT-----SDQVVLASRFWAYHKPLKNSQDYTEALRAA--RELAANITADLRKP-GTD 1087
Db 320 AVTASNSVISGGPFLVLTQTA-----PV-----SGLAAAFPTALAVLTITDVAVSPSGTG 369
Qy 1088 PAFEVFPVYITNVEY-----EQYLITLPEGLFMLSCLVPTFA-VSCILLGLDLRSLNLN 1142
```

```
Db 370 ---WIYGTSTRTLYAFASNGY---LPEIFLKIGKTKIPTYSLIAALIVGE----- 414
Qy 1143 LSIWMLIVDTVGFPMALWDISYNAVSLINLVSACMSV--EVSHITSEALSTKPTWLER 1200
Db 415 ----IFLLPFPFMYALVGFISSTATVLTYYIMGGIGLAVLRHAKELNRPFRV---PASVII 467
Qy 1201 AKREATISMGSA-----VPAGVAMTNLPGL-----VLGLAKAQLIQI----- 1237
Db 468 APIATLAAGLIVYWSSFAILFYVFTGIFL-GLPLFFIFYSNRIILGINKAYSIVVGVINLV 526
Qy 1238 -----PFRMLNLLITLLGLHLGLVPLPVILSVVGDV 1269
Db 527 IDLVMAFLFDFTSGLGAANNLFFGIYIVIVIAAMLIGDMFL---LKTVPDGV 576

RESULT 48
US-11-096-568A-2446
; Sequence 2446, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2446
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(726)
; OTHER INFORMATION: Ceres Seq. ID no. 15219050
; NAME/KEY: misc feature
; LOCATION: (628)..(628)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-2446

Query Match 1.5%; Score 100.5; DB 7; Length 726;
Best Local Similarity 21.8%; Pred. No. 11;
Matches 91; Conservative 47; Mismatches 162; Indels 117; Gaps 17;

Qy 46 SCSLMTLSNVCSLSTNPARKITGDHLILLOKICPLTYGTGNTQACCSAKQLVSLRASLSI 105
Db 220 SASPVSQDMVSNAPDIK-----PVVSGMTPELRTGPPGAVNLLNLSQVRQVMS 274
Qy 106 TKALLTRCPACSDNFNVLHCHNTCS-----PNQSLFI---NVTRVAQLGA----- 147
Db 275 SAALAGMASSVQSASAVAMHSMNISTGMATSLPPSQTVFTSQOQITSMAGSALMSAQ 334
Qy 148 -QQLPAVAYEAFYQHSFABQSYDSCSRVPAATLAV-----GTMCGVYSGALCNAQR 201
Db 335 TQSGPGP-----NNAFSPQTSN-----VASNLGVSPQMGNQNSHSGAMMQGGI 380
Qy 202 WLN-----FQDGTGNGLAPLDITFHLLPE--GQAVSGGIQPLNEGVARCNEQSDDDVA 252
Db 381 SMNQNMMSGLGQNVSSSGTGGMMPT-----PGVQQAQSGIQQLGG-----S 422
Qy 253 TCSCQDCAASCFAIARFQALDSTF-----YLGQMPGSLVLIILICSFVAVVTILLVGF 305
Db 423 NSSAPNQLSQSSGAMQTSQSKYVWEGNLSGQSQGQPVLI-----TRLGY 471
Qy 306 RVAPARD-----KSKMVPKGTSLSDKLSE---STHTLLGQFFQCGWT 346
Db 472 RSASASDLAANWPPTMQIVRLISQDHMNNKQTVGKADFLVFRAMSOHGLQQLQDKKLC 531
Qy 347 WVASWPLTILVLSVIPVWALAAGLVFTELTDDVELW-----SAPNSQARSEKAFHDQ 399
Db 532 AVIQLFSQTLILLSVDKACELIGMLF-----PGDMVVFQPIPNQOQOQQLHQ 582
```

RESULT 49
US-11-096-568A-2445
; Sequence 2445, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2445
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(727)
; OTHER INFORMATION: Ceres Seq. ID no. 15219049
; NAME/KEY: misc feature
; LOCATION: (629)..(629)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-2445

Query Match 1.5%; Score 100.5; DB 7; Length 727;
Best Local Similarity 21.8%; Pred. No. 11;
Matches 91; Conservative 47; Mismatches 162; Indels 117; Gaps 17;
QY 46 SGLMTLSNVCLSNTPARKITGDHLLILQKICPRLYTGPTNTQACCSAKQLVLEASLSI 105
Db 221 SASPVSDMVSNAPDIK-----PVVSGMTPLRTGPPGGANVLLNLSQVRQVMS 275
QY 106 TKALLTRCPACDNFVNLHCHNTCS-----PNQSLFI-----NVTVAQLGA----- 147
Db 276 SAALAGASSVGQSAVAMHMSNMISTGMTSLPPSTVSTGQGITSMAGSGALMGSAQ 335
QY 148 -GOLPAVAYEAFYQHSFARQSDVSCSRVRVPAATLAV-----GTWCGVYGSGALCNAQR 201
Db 336 TQSGPGP-----NNAFSPQTTN-----VASNLGVSQPMQGNQSHSGAMMQGGI 381
QY 202 WLN-----FQDGTGNGLAPLDITFHLLEP--GQAVSGIQPLNEGVARCNSQGDVVA 252
Db 382 SMNQNMMSGLGQGVNVSSTGGMPT-----PGVGQAQSGIQQLGG-----S 423
QY 253 TCSCDCAASCAPAIAPQALDSTF-----YLGQMPGSLVLIILCSVFAVVVILLVGF 305
Db 424 NSSAPNQLSQPSSGAMQTSQKYVWEGNLSGQSGQPVLII-----TRLEGY 472
QY 306 RVAPARD-----KSKMVDPKGTSLSDKLSF---STHTLLGQFFQGWGT 346
Db 473 RSASASDSLAANWPTMQIVRLISQDHMNNKQYVGKADFLVFRAMSHQGLQQLDKKLC 532
QY 347 WVASWPLTLLVSVIPVVALAAGLVTELTTPDVELW-----SAPNSQARSEKAFHQ 399
Db 533 AVIQLPSQTLILLSVSDKACRLIGMLP-----PGDMVVPKQIPNQQQQQQQQLHQ 583

RESULT 50
US-11-096-568A-2444
; Sequence 2444, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 2444
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(836)
; OTHER INFORMATION: Ceres Seq. ID no. 15219048
; NAME/KEY: misc feature
; LOCATION: (738)..(738)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-2444

Query Match 1.5%; Score 100.5; DB 7; Length 836;
Best Local Similarity 21.8%; Pred. No. 13;
Matches 91; Conservative 47; Mismatches 162; Indels 117; Gaps 17;
QY 46 SGLMTLSNVCLSNTPARKITGDHLLILQKICPRLYTGPTNTQACCSAKQLVLEASLSI 105
Db 330 SASPVSDMVSNAPDIK-----PVVSGMTPLRTGPPGGANVLLNLSQVRQVMS 384
QY 106 TKALLTRCPACDNFVNLHCHNTCS-----PNQSLFI-----NVTVAQLGA----- 147
Db 385 SAALAGASSVGQSAVAMHMSNMISTGMTSLPPSTVSTGQGITSMAGSGALMGSAQ 444
QY 148 -GOLPAVAYEAFYQHSFARQSDVSCSRVRVPAATLAV-----GTWCGVYGSGALCNAQR 201
Db 445 TQSGPGP-----NNAFSPQTTN-----VASNLGVSQPMQGNQSHSGAMMQGGI 490
QY 202 WLN-----FQDGTGNGLAPLDITFHLLEP--GQAVSGIQPLNEGVARCNSQGDVVA 252
Db 491 SMNQNMMSGLGQGVNVSSTGGMPT-----PGVGQAQSGIQQLGG-----S 532
QY 253 TCSCDCAASCAPAIAPQALDSTF-----YLGQMPGSLVLIILCSVFAVVVILLVGF 305
Db 533 NSSAPNQLSQPSSGAMQTSQKYVWEGNLSGQSGQPVLII-----TRLEGY 581
QY 306 RVAPARD-----KSKMVDPKGTSLSDKLSF---STHTLLGQFFQGWGT 346
Db 582 RSASASDSLAANWPTMQIVRLISQDHMNNKQYVGKADFLVFRAMSHQGLQQLDKKLC 641
QY 347 WVASWPLTLLVSVIPVVALAAGLVTELTTPDVELW-----SAPNSQARSEKAFHQ 399
Db 642 AVIQLPSQTLILLSVSDKACRLIGMLP-----PGDMVVPKQIPNQQQQQQQQLHQ 692

RESULT 51
US-11-087-099-4795
; Sequence 4795, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4795
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Oryza sativa
US-11-087-099-4795

Query Match 1.4%; Score 100; DB 7; Length 279;
Best Local Similarity 24.3%; Pred. No. 3.1;
Matches 46; Conservative 31; Mismatches 68; Indels 44; Gaps 9;
QY 1118 LCLVPTFAVSCLL--LGLDLRGLNLLISVIMLVDTVTFGMALMDSYNVSLINLVSIV 1175
Db 25 LCIIIP---VLCILEILFDKVRYSRNTKLSILVLVVG--VAVCTVTDVSVNSKGLLAIV 80
QY 1176 ---GMSVEFVSHITRSPAIS-----TKPT-----WLERAKETISMGSA 1211

```

Db      81   WSTALQQYVHHLQKKYSIGSFNLLGHTAPAAQAASLLILGPFDVFWLTNRREVDFTNYTTI 140
        1212 VPAGVAM-----TNLPGLVLGLAKAQLIQIPFFRLNLATLTLGLL---HGL---VF 1257
Db      141   VTFPIVLSCTIAVGNLNSOFICIGRTAVSPQVLGHMKTILVLTLGLFFPGKEGLNFHV 200
        1258 LPVILSYVG 1266
Db      201   LGMMLAVIG 209

RESULT 52
US-10-204-252-10
; Sequence 10, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.03340Z
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 3389
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: /Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-10

Query Match          1.4%; Score 100; DB 6; Length 3389;
Best Local Similarity 21.0%; Pred. No. 1.1e+02;
Matches       74; Conservative 49; Mismatches 127; Indels 102; Gaps 17;

Qy      762   ALTPMP-AVRTFALTSGLAVIDL--FLLOMSAFVALLSLDSKROEASRLDVCCKVPQEL 818
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      2135   ALESLPTLELLLTLTATVTGGIFLFMSA-----RGIGWTLGMCCIIT---- 2181
        |||::|||::|||::|||::|||::|||::|||::|||::|||::

Qy      819   PPPQGSGELLGFQKAYAPFLHWITRQVVLLFLALFGVSLSYMSCHISVGLDQELALP 878
        |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      2182   -----ASILW-----YAQIQPHWTAASIIIEFFLIIV-----LLIPEPEKQRT 2220
        |||::|||::|||::|||::|||::|||::|||::|||::|||::

Qy      879   KDSVLLDYFLPLRYEVGAPVYFVITLGVNFSSEAGMAINACSSAGCNMFSTQIKQ--- 935
        |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      2221   QDNQL-----TVVIAI-----LTVVAATWANEMGLETKKDL 2254
        |||::|||::|||::|||::|||::|||::|||::|||::|||::

Qy      936   ----YATEPFQSYLAI---PASSSWDDFDIMLTSSCCRLYTISGNKDQFCFSTVN-SL 987
        |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      2255   GLGSIAITQPESNILDLIDLPASAW-----TLVAVATTFTVPMLRHSIENSSNVSL 2306
        |||::|||::|||::|||::|||::|||::|||::|||::|||::

Qy      988   NCLKNCMSITWGSVRPSVEGFHKYLPWLNDPRNIKCPKGGLAAYS--TSVNLTSDGQVL 1045
        |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      2307   TAIANQATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYSQVNPILTALFIL 2355
        |||::|||::|||::|||::|||::|||::|||::|||::|||::

Qy      1046   ASRFWAYHKPKNSQDYTEALRAARELANITA-----DLRKVPCTDPAFE 1091
        |||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      2356   VAHYAITGPGLQAKATREAOKRAAGAIMKNPTVDGITVIDLDPIP-YDKPFE 2406
        |||::|||::|||::|||::|||::|||::|||::|||::|||::

RESULT 53
US-10-204-252-6
; Sequence 6, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the

```

US-10-204-252-8

Query Match 1.4%; Score 100; DB 6; Length 3391;
Best Local Similarity 21.0%; Pred. No. 1.1e+02;
Matches 74; Conservative 49; Mismatches 127; Indels 102; Gaps 17;
QY 762 ALTPMP-AVRTFALTSGLAVID--FLLQMSAFVALLSLDSKQASRLDVCCVKPQEL 818
DB 2137 ALSELPETLETLTLLATVTGGIFLFLMSA-----RGIGKWTLMGCCIIIT--- 2183
QY 819 PPPQGGEGLLGFFQKAYAPFLLHWTTRGVLLFLALFGVSLYSNMCHISVGLDQELALP 878
DB 2184 -----ASILLW-----YAIQIOPHWIAASIIIEFFLIV-----LLIPEPEKQRTP 2222
QY 879 KDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSSEAGMNAICSSAGCNPFSTQKIQ--- 935
DB 2223 QDNQL-----TVVIAI-----LTVVAATWANEMGFLEKTKDOL 2256
QY 936 -----YATEPPEQSYLAI-----PASSWVDDFDIMLTPSSCCRLYISGNKDKFCPSTVN-SL 987
DB 2257 GLGSIATQOPESNILDIDLRPASW-----TLVAVATTFVTPLMRHSIENSSVNVSL 2308
QY 988 NCLKNCMSITMGSVRSVQFHKYLPWFLNDRPNKICPKGGLAAYS--TSVNLTSQGQVL 1045
DB 2309 TAIANOATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYSQVNPITLTAALFLL 2357
QY 1046 ASRFMAYHKPLKNSQDYTEALRAARELANITA-----DLRKVPGTDDPAPE 1091
DB 2358 VAHYAIIIGFGLQAKATREAKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2408

RESULT 55

US-10-204-252-12
; Sequence 12, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.033402
; CURRENT APPLICATION NUMBER: US/10/204,252
; PRIOR FILING DATE: 2002-02-16
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-12

Query Match 1.4%; Score 100; DB 6; Length 3391;
Best Local Similarity 21.0%; Pred. No. 1.1e+02;
Matches 74; Conservative 49; Mismatches 127; Indels 102; Gaps 17;
QY 762 ALTPMP-AVRTFALTSGLAVID--FLLQMSAFVALLSLDSKQASRLDVCCVKPQEL 818
DB 2137 ALSELPETLETLTLLATVTGGIFLFLMSA-----RGIGKWTLMGCCIIIT--- 2183
QY 819 PPPQGGEGLLGFFQKAYAPFLLHWTTRGVLLFLALFGVSLYSNMCHISVGLDQELALP 878
DB 2184 -----ASILLW-----YAIQIOPHWIAASIIIEFFLIV-----LLIPEPEKQRTP 2222
QY 879 KDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSSEAGMNAICSSAGCNPFSTQKIQ--- 935
DB 2223 QDNQL-----TVVIAI-----LTVVAATWANEMGFLEKTKDOL 2256

QY 936 -----YATEPPEQSYLAI-----PASSWVDDFDIMLTPSSCCRLYISGNKDKFCPSTVN-SL 987
DB 2257 GLGSIATQOPESNILDIDLRPASW-----TLVAVATTFVTPLMRHSIENSSVNVSL 2308
QY 988 NCLKNCMSITMGSVRSVQFHKYLPWFLNDRPNKICPKGGLAAYS--TSVNLTSQGQVL 1045
DB 2309 TAIANOATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYSQVNPITLTAALFLL 2357
QY 1046 ASRFMAYHKPLKNSQDYTEALRAARELANITA-----DLRKVPGTDDPAPE 1091
DB 2358 VAHYAIIIGFGLQAKATREAKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2408

RESULT 56

US-10-204-252-16
; Sequence 16, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.033402
; CURRENT APPLICATION NUMBER: US/10/204,252
; PRIOR FILING DATE: 2002-02-16
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-16

Query Match 1.4%; Score 100; DB 6; Length 3391;
Best Local Similarity 21.0%; Pred. No. 1.1e+02;
Matches 74; Conservative 49; Mismatches 127; Indels 102; Gaps 17;
QY 762 ALTPMP-AVRTFALTSGLAVID--FLLQMSAFVALLSLDSKQASRLDVCCVKPQEL 818
DB 2137 ALSELPETLETLTLLATVTGGIFLFLMSA-----RGIGKWTLMGCCIIIT--- 2183
QY 819 PPPQGGEGLLGFFQKAYAPFLLHWTTRGVLLFLALFGVSLYSNMCHISVGLDQELALP 878
DB 2184 -----ASILLW-----YAIQIOPHWIAASIIIEFFLIV-----LLIPEPEKQRTP 2222
QY 879 KDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSSEAGMNAICSSAGCNPFSTQKIQ--- 935
DB 2223 QDNQL-----TVVIAI-----LTVVAATWANEMGFLEKTKDOL 2256
QY 936 -----YATEPPEQSYLAI-----PASSWVDDFDIMLTPSSCCRLYISGNKDKFCPSTVN-SL 987
DB 2257 GLGSIATQOPESNILDIDLRPASW-----TLVAVATTFVTPLMRHSIENSSVNVSL 2308
QY 988 NCLKNCMSITMGSVRSVQFHKYLPWFLNDRPNKICPKGGLAAYS--TSVNLTSQGQVL 1045
DB 2309 TAIANOATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYSQVNPITLTAALFLL 2357
QY 1046 ASRFMAYHKPLKNSQDYTEALRAARELANITA-----DLRKVPGTDDPAPE 1091
DB 2358 VAHYAIIIGFGLQAKATREAKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2408

RESULT 57

US-10-204-252-28
; Sequence 28, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:

Query Match 1.4%; Score 99.5; DB 6; Length 577;
Best Local Similarity 18.4%; Pred. No. 9.5;
Matches 69; Conservative 48; Mismatches 116; Indels 141; Gaps 18;

QY 865 CHI--SVGLDDELA--LPKDSYLL-----DYFLFLNRYFEVGPV-----YFV 903
DB 93 CHIEKAGTDDQNKKEYCSKEGHILIECGAPRNGKRSKSLSTAVTILTGSLVTVAEQPP 152
QY 904 TITGYNFSS-----EAGMNAICSSAGCANNFSFTQIKQVATER--PEQSY 945
DB 153 VTVVRNFRGLAELLKVGSKQKQEDKWTAVHVIIVGPGCG-----KQWARNFAERDYY 206
QY 946 LAIPASW-----VDDPIDWLTSSCCRLVISGPKDKFCPCSTVNSLNCNKCM 994
DB 207 WKPSRNKKWDGVBGEVWVLDLDFYGLPWLDLRLC-----DRY-PLTVET----- 251
QY 995 SITMGSVRSVGFQHKYLPWFLNDRNPKCPKGLAAY-STSVNLTSD-----QCVLA 1046
DB 252 -----KGGTVPFILARSILITISNQAPQEWSSSTAV 280
QY 1047 SRFMAYHKPKNSQDYTEALRAARELANITADLRKVPDPAFEVFPYITNVFY----- 1102
DB 281 PAVEALYRITTLQFWKTAGEQSTVEPEG-----RFEAVDPFCALFPYKINVFVITS 334
QY 1103 -----BOYLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLL--NLSI 1145
DB 335 WFLFLFIRVQDKFSLYNSQPHIILGCGCILERIAQCV--LDIGVIMEPQLVGF 391
QY 1146 VMLVDVTGFMALW 1159
DB 392 I-IWLEPINCLVLW 404

RESULT 60
US-10-506-454-1056
; Sequence 1056, Application US/10506454
; Publication No. US2006068386A1
; GENERAL INFORMATION:
; APPLICANT: Mezhevaya, Alexi I
; APPLICANT: Polushin, Nikolai N
; APPLICANT: Shcherbinina, Olga V
; APPLICANT: Shakhova, Vera V
; APPLICANT: Malykh, Andrei G
; APPLICANT: Kozvavkin, Sergei A
; TITLE OF INVENTION: The Complete Genome and Protein Sequences of the Hyperthermophile
; TITLE OF INVENTION: Methanopyrus kandleri AV19 and Monophyly of Archaeal Methanogens
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: FID001
; CURRENT APPLICATION NUMBER: US/10/506,454
; PRIOR FILING DATE: 2004-08-31
; PRIOR FILING DATE: 2003-03-04
; PRIOR FILING DATE: 2003-03-04
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1722
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1056
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Methanopyrus kandleri
US-10-506-454-1056

Query Match 1.4%; Score 99; DB 6; Length 716;
Best Local Similarity 17.2%; Pred. No. 14;
Matches 84; Conservative 88; Mismatches 159; Indels 156; Gaps 19;

QY 846 RGVLILLFALFGVSLYSMCHISVGLDDELAIPKDSYLLDYFLFLNRYFEVGPVYFVTT 905
DB 4 RTAILUSFLALTAVMAYGALHNVEDVDQTKYLPD----- 37
QY 906 LGYNFSSEAGMNAICSSAGCANNFSFTQIKQY--ATEFPQSYL-----AIPASSWVDDFI 958

DB 38 ----RFESMKQWVVERELGTSTKTLIVIEADDDVTRKPVLDYMRRIEDRLRSKPYVEN-- 92
QY 959 DMLTSSCCRLVISG--PNKDKFCPCSTVNSLNCNKCMSTMGSVRSVSPQFHKYLPWFLN 1017
DB 93 -----VRGAPDVLRESPPA-----AVTMCGLRPLMSEMER----- 125
QY 1018 DRPNIKCPKGLAAYSTSVNLTSDQVTLASRFMAVHKPLKNSQDYTEALRAARELANIT 1077
DB 126 -----SKEMFVSKHKVAIRV-----GLKSDADYRK-----VV 154
QY 1078 ADLRKVPDPAFEVFPYITN---VFYEQYLTILPEGLFMLSCLVCLVPTFAVSCLL----- 1130
DB 155 PDVRSLSERDKPKSVKPADVTGSPAINYDFYRGLKD---LVTVTALVSAVAALVYVDF 211
QY 1131 -----LGLDLRSGLLNLISYVIMLVDTVGFMAWLDISYNAVSLNLINLAVSGMSVEFV- 1182
DB 212 RRWAPVGLTI---ILSAVANVGLIMYWLGFAPFY-----ATVLTVMVGVDYVI 263
QY 1183 ---SHITRSFAISTKP-----TWLRAKEATISMG---SAPVAGVAMTNLPGILVLGLA 1230
DB 264 FTLTRFQBEYDIKGRAGKGEALLTAVRAGRAVLITGLTASAGFAALALSEFRMVSBIG-- 321
QY 1231 KAQLQIFFRNLNLTILGLHLGLVFLPVILSYGDPVNPALALEOKRAEBAVAVMVA 1290
DB 322 -----LGIVAGILTAVALTLLVPLSLLOSIPIGRKSEKSEKSEKSKVVL 363
QY 1291 SCP-NHP 1296
DB 364 SIPVRHP 370

RESULT 61
US-10-204-252-14
; Sequence 14, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.0334U2
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3391
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-14

Query Match 1.4%; Score 98; DB 6; Length 3391;
Best Local Similarity 19.8%; Pred. No. 1.6e+02;
Matches 69; Conservative 52; Mismatches 132; Indels 96; Gaps 15;

QY 762 ALTPMPAVRTFALTSGLAVIDPLQLQMSAFVALLSDSKRQBSRLDVCCKVPQBLPPP 821
DB 2137 AUSELPE-----TLETLLLLTLATVTGIGIFLMSGRIGIKMTLGMCCIIT----- 2183
QY 822 GQGEGLLGGFFQKAYAPFLHMTIRGVILLFLFALFGVSLYSMCHISVGLDDELAIPKDS 881
DB 2184 ---ASILLW-----YAOIQPHWIAASIIILEFLIV-----LLIPEPEKQRTFQDN 2225
QY 882 YLLDYFLFLNRYFEVGPVYFVTTLGYNFSSEAGMNAICSSAGCANNFSFTQIKQY----- 935
DB 2226 QL-----TYVVIAT-----LTVVAATWANEMGFLEKTKDGLG 2259

```
QY 936 -YATFPESYLAI---PASSWDDFDLWTFSSCCRLIYISPNKDKFCPTVN-SINCL 990
Db 2260 SIATQOPESNILDLRPASW-----TLYAVATFTVPMULRHSIENSNNVSLTAI 2311
QY 991 KNCMSITWGSVRPSVEQFHKYLPWFLNDRPNIKCPKGGLAAYS--TSVNLTSDDGOVLASR 1048
Db 2312 ANQATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYSVQNPITLTAALFLVAH 2360
QY 1049 FMAHYHKLKNSQDYTEALRAARELANITA-----DLRKVPGTDPAFE 1091
Db 2361 YAIIGPGLQAKATREAQKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2408

RESULT 62
US-10-204-252-18
; Sequence 18, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.0334U2
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3402
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-18
```

```
Query Match 1.4%; Score 98; DB 6; Length 3402;
Best Local Similarity 19.8%; Pred. No. 1.6e+02;
Matches 69; Conservative 52; Mismatches 132; Indels 96; Gaps 15;

QY 762 ALTPMPAVRTFALTGLAVILDFLQMSAFVALLSDSKRQESRLDVCCKVQBLPPP 821
Db 2148 ELSEPE-----TLETLLLTLLATVTGGIFLMSGRGIGRWTLGMCIIIT----- 2194
QY 822 GQEGILLGFFQKAYAPFLHWTIRGVLLFLALFGVSLYSNCHISVGLDQELAPKDS 881
Db 2195 ---ASILLW-----YAQIQPHWIAASIIIEFFLIV-----LLIPEPEKQRTPDN 2236
QY 882 YLLDYFLFLNRYPEVGAPVVFVTTLGVNFSSEAGMNAICSSAGCANNFSTQKIQ----- 935
Db 2237 QL-----TVVTIAI-----LTVVAATWANEMGFLEKTKDGLG 2270
QY 936 -YATFPESYLAI---PASSWDDFDLWTFSSCCRLIYISGNKDKFCPTVN-SINCL 990
Db 2271 SIATQOPESNILDLRPASW-----TLYAVATFTVPMULRHSIENSNNVSLTAI 2322
QY 991 KNCMSITWGSVRPSVEQFHKYLPWFLNDRPNIKCPKGGLAAYS--TSVNLTSDDGOVLASR 1048
Db 2323 ANQATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYSVQNPITLTAALFLVAH 2371
QY 1049 FMAHYHKLKNSQDYTEALRAARELANITA-----DLRKVPGTDPAFE 1091
Db 2372 YAIIGPGLQAKATREAQKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2419
```

```
RESULT 63
US-11-087-099-4100
; Sequence 4100, Application US/11087099
; Publication No. US20060041961A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4100
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(324)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-4100
```

```
Query Match 1.4%; Score 97; DB 7; Length 324;
Best Local Similarity 22.0%; Pred. No. 6.6;
Matches 53; Conservative 38; Mismatches 80; Indels 70; Gaps 12;

QY 1088 PAPEVFPYTIINVFEQYLITLPEGLFML-----SLCLVPTFAVSCLL--LGLDL 1135
Db 54 PSYLPPELVKVFVFAN-LSIVGMNVSLMWNVSFGYQIAKLCIIP---VLCLEILFDKV 109
QY 1136 RSGLLNLLSIVMILVDVTGFMALWDISYNVSLINLVSAV---GMSVEFVSHITRSFAIS 1192
Db 110 RYSRNTKLSIVLVVG-VAVCTVTDVSVNSKGLLAAVIAVWSTALQOHVHVHLQKKYSLG 168
QY 1193 -----TKPT-----WLRAKEAT-----ISMGSVA 1212
Db 169 SFNLLGHTAPAAQASLLILGPPVDFLTNRVDITFNYYTITVTLFHLNLFQFVILSCTI 228
QY 1213 FAGVAMTNLPGILVGLAKAQLIQIFPFLNLLITLGLL---HGL---VFLPVILSV 1265
Db 229 AVG---TNLSQFICIGRFTAVSFQVLGHMKTILTLGLFPGKGLFHVHVLGMLLAVI 285
QY 1266 G 1266
Db 286 G 286
```

```
RESULT 64
US-11-087-099-7384
; Sequence 7384, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7384
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(336)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-7384
```

```
Query Match 1.4%; Score 96.5; DB 7; Length 336;
Best Local Similarity 22.7%; Pred. No. 7.6;
Matches 52; Conservative 38; Mismatches 84; Indels 55; Gaps 11;

QY 1088 PAPEVFPYTIINVFEQYLITLPEGLFML-----SLCLVPTFAVSCLL--LGLDL 1135
Db 71 PSYLPPELVKVFVFAN-LSIVGMNVSLMWNVSFGYQIAKLCIIP---VLCLEILFDKV 126
QY 1136 RSGLLNLLSIVMILVDVTGFMALWDISYNVSLINLVSAV---GMSVEFVSHITRSFAIS 1192
```


Db		127	RYSDTFESIMVVLVG--VANVCTVDVSVNSQGLIAALIANWSTALQOHVYHLLKQKYSLG	185
QY		1193	-----TKPT-----WLERAKGATISMGSAVFAGVAM-----TNLPGI	1224
Db		186	SPNLLGHTAPAAQASLLVLGFDVFWLTKRVDFTFNVTIAVTFPIILISCTIAVGNLSQF	245
QY		1225	LVLGLAKAQAIQIQFFPFRNLNLTLLGLL-----HGLVF---LPVTLISVVG	1266
Db		246	ICIGRFTAVSFGVLGHMKTVLVLTLGLFFFGKEGLNPHVAIGMLTAVIG	294

RESULT 65

```

US-10-793-626-2928
; Sequence 2928, Application US/10793626
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ.ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2928
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2928

```

Query Match	1.4%	Score	96.5;	DB	6;	Length	465;
Best Local Similarity	23.14;	Pred.	No. 12;				
Matches	93;	Conservative	53;	Mismatches	131;	Indels	125;
Gaps	26;						
QY	574	FSLNNTYP-AGDPRLQAQKLEWEEAPLEBMRAPFORRMAGMFQVTF					TAER 619
Db	10	FDIHKMKNKGIPLAVQEKMLRNF					TAER 619
QY	620	SLEDEINRTTAEDLPATPATSYIVIFLYISIALSGYSWSRWVDSKATLGLGGVAVLGA					679
Db	60	LLKKEEIGLTLE					679
QY	680	V-NAAMGF-PSYLGRSSLILOVWPPLVSVGDNIFFIV					733
Db	112	ITVLIWGFVUSYFG					163
QY	734	VHIG					787
Db	164	RYLGFWNTSHNIGAITAGGVALWG					209
QY	788	MSAFVALLSDSKROEASRLDVCCKVQBEL					833
Db	210	IGIVTLFIKQDB					254
QY	834	KAYAPFLHWHITRGVWLLFLFLAGVSLYSWMCHISVGLDQELAPKDSYLLDYFLFLNR					892
Db	255	KKY					892
QY	893	YPRVGAPVYVVTTLGNVFSSE					925
Db	299	DAVNTIFYFEIGA					336

RESULT 66

```

US-11-087-099-2386
; Sequence 2386, Application US/11087099
; Publication No. US20060041961A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
```

Db 712 L-----WLKRHNSRLKFIHWPVPFNGPSQ 736

RESULT 67

US-10-858-730-118
; Sequence 118, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 294
; TYPE: PRF
; ORGANISM: Lactobacillus plantarum
US-10-858-730-118

Query Match 1.4%; Score 96; DB 6; Length 294;

Best Local Similarity 22.5%; Pred. No. 6.9;
Matches 49; Conservative 42; Mismatches 83; Indels 44; Gaps 9;

QY 1085 GTDPAFEPVPTITNVFYEQYITLPEGLFMLSCLVPTFAVSCLLGL--DLRGLNL 1142
DB 76 GTASGF-----TNTFYIQALKLIP--VAVAAVLMQAVWISTLLGAVIHRRPSRLQV 126
QY 1143 LSVMLVDTVGFEMALWDISYNAVSLINLVSAVGSVEFVS-----HITRFAISTK 1194
DB 127 VSVLVLTGTLAAGLFPT-----QALSPWGLMLSFLAACSVACTMQFTASLGNL 179
QY 1195 P---TWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLGL 1251
DB 180 PLSKTWL-----LCLGA--FLIIAIVSPQLVTPATPTATV-----GWGLIALFSM 224
QY 1252 LHGLVFLPVLSVGVDPVNPALALEQKRAEEAAVAVV 1289
DB 225 VPELVMSYLFMPYLEGIGIPILSLSLELPASIVVAVFLL 262

RESULT 68

US-11-087-099-4490
; Sequence 4490, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4490
; LENGTH: 479
; TYPE: PRF
; ORGANISM: Escherichia coli K12
US-11-087-099-4490

Query Match 1.4%; Score 96; DB 7; Length 479;
Best Local Similarity 18.7%; Pred. No. 14;
Matches 112; Conservative 74; Mismatches 175; Indels 238; Gaps 26;

QY 721 YQRLPRRPGEPREVIHIGRALG-----RVAPSMLLCSLSBAICFFLGALTMPAVRTF 772
DB 6 HDRKPYTAGEAAIMAINSPINIAAOPGKTRLRKSL---KLQVVMGLAYLTPMTVFDTF 62
QY 773 ALTSGLA---VILDFLLQMSA--FVALLSLDSKRQEASRLDVCCVCPQEPQOQGRGL 827
DB 63 GIVSGISDGHVPASILLALAGVLTALSYGKLVQ-----FPEAGSA--- 104
QY 828 LLGFFQKAYAP---FLHFWITRGVVLFLFLALFGVLSYMSCHISVGLDQELALPKDSYLL 884
DB 105 -YTYAQKSINPHVGFVGM-----SSLL 126
QY 885 DYFLFLNRYFVGAPVFTVTLGNFSEAGMNAICSSAGCNCNFSTOKIQIYATFPPQS 944
DB 127 DY-LFL-----PM-----INVLLA-----KIYLSALFPE-- 149
QY 945 YLAIPASSWDDFDLTPSSCCRLYISGPNKDKPCSTVNSLNLCKNC-----MSI 996
DB 150 ---VPPWVWVTFVAILTAAN-----LKSVMNVANFTLFLVLQISL 188
QY 997 TMGSVRPSVEQFHK-----YLPWFNLDRPNIKCPKGLAAVSTSVNLTSDGQVLASR 1048
DB 189 MVVIFLVQGLHKGEGVGTWSLQPFISENAHL-----IPIITGATIVCFS 235
QY 1049 FMAYHKPLKNSQDYTEALRAARELANITA-----DLRKEVGTDBA 1089
DB 236 FLGFDVAVTTLSEETPDARVIPK-AIFLTAVYGVGFIAAASFPMQLFPDPSRFXDPDAA 294
QY 1090 FEVPYITITNVFYEQYITLPE-----GLFMLSCLVPTFAVSCLLGLDLRSGLNLL 1143
DB 295 -----LPEIALYVGKLFQSIPLCTTP--VNTLASGLASHASVSRLL 334
QY 1144 STV-----MILVDTVGFMALWDISYNAVSLIN---LVSAVG 1176
DB 335 YVMGRDNPVPERVGVHPKWRTPALNVIMGVIALSALFFDLVTATALNFGALVAFTF 394
QY 1177 MSVEFVSHITSFAISTKPTWLERAKEATISMGSAVAGVAMTNLPGI-LVLGLAKAQL 1234
DB 395 VNLSVFNFHFRKKGNN--KSWKDHFYLLMPLVGLVGLVWVNLSTSLTGLVWASL 451
RESULT 69
US-11-087-099-10324
; Sequence 10324, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10324
; LENGTH: 479
; TYPE: PRF
; ORGANISM: Escherichia coli O157:H7
US-11-087-099-10324
Query Match 1.4%; Score 96; DB 7; Length 479;
Best Local Similarity 18.7%; Pred. No. 14;
Matches 112; Conservative 74; Mismatches 175; Indels 238; Gaps 26;
QY 721 YQRLPRRPGEPREVIHIGRALG-----RVAPSMLLCSLSBAICFFLGALTMPAVRTF 772
DB 6 HDRKPYTAGEAAIMAINSPINIAAOPGKTRLRKSL---KLQVVMGLAYLTPMTVFDTF 62
QY 773 ALTSGLA---VILDFLLQMSA--FVALLSLDSKRQEASRLDVCCVCPQEPQOQGRGL 827
DB 63 GIVSGISDGHVPASILLALAGVLTALSYGKLVQ-----FPEAGSA--- 104


```
Db 165 PELRQLQALLQRPHQNTTGTTRPCWGSTHPRKASONEBAPLRQEQNQLRQKLPQGA 224
Qy 493 -----YFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTFKDGTALALS 541
Db 225 EAGTVLAGELGFLAOPGAFRLRPNVVLGSLTEVLSRFPFCLLLGP----- 272
Qy 542 CWADYGAPVFPPLAIGGYKGDYS--AEALIMTFSLNNYPAGDPRLAQAKLME----- 593
Db 273 CML-----GKGHEWGRAAAVLLS-----DPQFQ-----NSVRRAS 303
Qy 594 -----EAPLEWRAP-----QRMMAGMFQVTFTA 617
Db 304 NLHLLAALDALEEVTVLPGRWDPTARIPPKCLPSQHKELPSQREIRGPAVRLUT- 362
Qy 618 ERSLED-----EINRT-----TAEDLPIFATSYI-----VIFLY 646
Db 363 --SAEDRRHGHGPHAHSPQLQRTGRLFGGLIQDVRKVPWYPSDFDLALHLCFSVLYIY 420
Qy 647 ISLALGSYSWSRVMVDSKATILGLGSV-AVVLGAVMAAGRFSYLG---IRSSLVILQV 701
Db 421 LATVTNAI--TFGGLGD--ATDGAQGVLESFGLTAVAGAAFCLMAGQPLTILSSGTVLV 477
Qy 702 VPEFLVSGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEALCFPLG 761
Db 478 FERLLFSFSD-----YSLDY--LPR-----LWVGIVWA-----TFCLVLV 512
Qy 762 ALTPMPAVRTFA--LTSGLAVILDFLQMSAFVALLSLDSK---RQEARLDDVCCVCPQ 816
Db 513 ATEASVLVRYFRFTBEGFCALISIFIYDAVGKMLNLTHTYPIQKPGSSAYGCILC---- 568
Qy 817 ELPPPGQGGELLGFPQKAYAPFLHMI-----TRGVLLLFALFGVSLYSM----- 864
Db 569 QYPPGPGNES-----QWTRPKDRDDIVSMDLGLINASLLPPPECTROG 613
Qy 865 -----CHISVGLDQELALPKDSYLLDYFLNRYFE-----VGAPYFVFTILGYNPS 911
Db 614 GHPRGPGCH-----TVP-DIAFFSLLFLTSPFAMALCKVTSRFPSPVVRKGLS 663
Qy 912 SEAGMNAICSSAGCNF--SFTQKIQYATEP-----PEQSYLAIP--ASSWDDDFDILWLP 963
Db 664 DFSSVLAILLGCGDLAFLGLATPKLMVPREFKPTLFGRGWLVSPPGANPMW-----WSVA 718
Qy 964 SSCRLYIS-----GPNKDFCPSTVNSLNCCLKNCMSITM 998
Db 719 AALPALLSLIPMDQOITAVILNRMEYRLQKAGFHLDFCVAVLMLL-----TSAL 771
Qy 999 GSVRPSVEQFHKYLPFLWFLNDPRNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKN 1058
Db 772 G-----LPWVVS-----ATVISLA----- 785
Qy 1059 SQDYTBALRAARELANITADLRKVPGTDPAP-----EVPYITVNVFEOYILT 1107
Db 786 --HMSLRRESRACA-----PGERPNFLGIREQRLTGLVWFLTGASIFLAPVLK 833
Qy 1108 ILPEGLFMLSCLAVPTFAVSCLLGLDLRSGLIANLLSI 1145
Db 834 FIP-----MP-----VLYGIFYMGVAALSSI 855

RESULT 72
US-11-097-125-2
; Sequence 2, Application US/11097125
; Publication No. US2005028600A1
; GENERAL INFORMATION:
; APPLICANT: HADLEY, GREGG ALLEN
; TITLE OF INVENTION: THERAPEUTIC BLOCKADE OF CD103 INTERACTIONS TO PREVENT
; TITLE OF INVENTION: CLINICAL RENAL ALLOGRAFT REJECTION
; FILE REFERENCE: 0402.003.0002
; CURRENT APPLICATION NUMBER: US/11/097,125
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/559,255
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 2
```

```
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Mus sp.
US-11-097-125-2

Query Match 1.4%; Score 96; DB 7; Length 1167;
Best Local Similarity 22.7%; Pred. No. 49;
Matches 96; Conservative 49; Mismatches 157; Indels 120; Gaps 19;

Qy 363 VVALAAGLVFTELTPVELMSPANSQARSKAFHQHGFPRFTNQVILFAPNRSSTRY 422
Db 300 MVLVTGDIFF--GDPLNLTTVINSPKMQGVWVFAIGVGDRFK-----NNNTYRE 346
Qy 423 DSLILG-PK-----NFSGILDLDLLELLELQERLHLQ-----VMSPEAQORNLSL 467
Db 347 LKLIASDPKEAHTFKVTNYS--LDGLLS--KLQQRIVHMEGTVDGLQYLAQTGFSA 401
Qy 468 QDICVAPLNPNTSLYDCCINSLLQYFONNRTLLL-----LTANQTLMGQTSQVDWK 519
Db 402 QILDKGVLLTGVAPNWSGGALLYSTQNGRCFLNQAKEDSRVQYSLGYSLAVLHK 461
Qy 520 DHFL-YCANAPLTFKOGTALALS-----CWADY---GAP 549
Db 462 AHGISYVAGAPRHKLGRGAVFELRKEDREDAFVRRIEGEQMSGYFSGVLCPVDDMDGTT 521
Qy 550 VF-----PFLAIGGYKGDY-----SEAEALIMTFSLNNYPAGDPRLAQAKLMEEAFLSEM 600
Db 522 DFLVAAFPFYHIRGEGRVYVYQPEQDA---SFLAHTLSGHPGLTNSR----- 568
Qy 601 RAFQRRMAGMQVTFTAERSLEDEINRTTABDLPIFA-----TSYIVIFLYISLA 650
Db 569 --FGFMAAV-----GDIQDKFTDVAIGAPLEGFGAGDGASGYSVYIYNGHS 614
Qy 651 LGSYSSSRVMDSKATILGLG--GVAVVLGAVMAAGFFSY-LGIRSSLIVILQVVPFLVL 707
Db 615 GGLYDSPSQOIRASSVASGLHYFGMSVSGGLDFNGDLADITVGSRDSAVVLSRFPVVDL 674
Qy 708 SV 709
Db 675 TV 676

RESULT 73
US-11-087-099-8709
; Sequence 8709, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8709
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(338)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-8709

Query Match 1.4%; Score 95.5; DB 7; Length 338;
Best Local Similarity 21.7%; Pred. No. 9.2;
Matches 50; Conservative 41; Mismatches 84; Indels 55; Gaps 11;

Qy 1088 BAFEVFPYITITNVFVEQYLTILPEGLFML-----SLCLVPTFAVSCIL--LGLD 1134
Db 75 PSYLPPELVKVFVFAN--LSIVGMNVSLMWSNVGXGYQIAKLCII--XPVLCTLETWTWK 131
```

```
QY 1135 LRSGLLNLSIVMLVDTVCFMAMLDISYNAVSLINLVSAV---GMSVFVSHITSPAI 1191
Db 132 VRYSDTKLSIVLVLVG-VAVCTVTDVSVNSKGLLAAVTAWSTALQQHYVHHLQKYSL 190
QY 1192 S-----TKPT-----WLERAKEATISMGSAVPAVAM-----TNLPG 1223
Db 191 GSFNLGHTAPAAQASULLILGPPVDFWLTNRVDTFTNYTITVTFVLSCTTAVGNTLSQ 250
QY 1224 ILVLGLAKAQLOIQIFFRNLNLTLLGLL-----HGL---VFLPVILSYVG 1266
Db 251 FICIGRFTAVSFOVLGHMKTILVLTGLFFGKEGLNFHVLGMLAVIG 300

RESULT 74
US-10-821-234-1027
; Sequence 1027, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1027
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1027

Query Match 1.4%; Score 95.5; DB 6; Length 686;
Best Local Similarity 18.8%; Pred. No. 25;
Matches 148; Conservative 87; Mismatches 254; Indels 297; Gaps 37;

QY 601 RAFORMAGMFQVTFABRSLEDEINRTTAEDLPFATS-----YI--VIFL 645
Db 36 RALEQOLA-----ADRHSEDMESGVSTAPLSPTSGTSGMSTFTSIMDYVVFVLL 87
QY 646 YISLALGSYSS---MSR-----VMVDSKATLGLGVAVVLGAVMAAMGFFSYLGI 696
Db 88 VLSLAIGLHACRGHRTVGLLMADRK---MGCLPVLSLTLATFQSAVAILGVPS-- 141
QY 697 VILQVVPVLVSVGADNIFVLEVQRLPRPGRPREVHIGRALGRVAPSMLLCSLEAI 756
Db 142 -----EIVRFGTQYWF-----GC 155
QY 757 CPELGLALTP---MPAVRTFALTSLGLAVI-----LDLFLQMSAP--VALLSL 797
Db 156 CYPLGLLIPAHIFIPVYFRLHTSAYEYLELRFNKTVRVCGTVTFIFQVITMGVVL 215
QY 798 DSKROEASRLDCCVCKPQELPPGQGGELLLGFFQKAYAPF---LHWTIRGVVLLIF 853
Db 216 SLALNAVTFDLWLSV-----LALGIUVCVTVYALGGLKAVITDVFQTLWF 262
QY 854 LALFGVSLYSMCHISVGLDQELALPK-----DSYLLDYFLFLNRYF---EVGAP 905
Db 263 LQLAVIIVGSAAKVG-GLGRVNAVASQHGRIISGFELDPDPFVRHTFTLAFGCVF 321
QY 906 LGYN-----PSSEGNMAICS-----SAGC-----NNPSTQK-----IQ 938
Db 322 YGNOAQVORYLSRTEKAAVLSCYAVFPFQVSLCVGCLIGLVMPAYQYQYPMISQ 381
QY 939 EPPEGSYLAIPASSVDDFDWL---TPSSCCRLXISGPNKDFCPTSTVNSLNCNCSI 996
Db 382 AAPDQFVL-----YFMDLLKGLPLGLFIACLPSG-----SLSTISS---AF 428
QY 997 TWGSVRPSVEQPHKYLPPWF--LNDRPNIKCPKGLAAVSTSVNLTSDGQVLASRF 1054
```

```
Db 429 TMEDL-----IRPNWPFSEARAIMLSRGLAFGYG-----LLCLGMAY-- 466
QY 1055 PLKNSQDYTEALRAARELAANITADLRKVPDPAFEVPPYTTITNVFYEQYLTPGL 1114
Db 467 ---ISQMGFPVQAALISIFGMVGGPLL-----GLP 493
QY 1115 MLSL---CLVPTFAVSCLLGLDLRLSGLNLLLSIVMLVDTVGFMALWDISYNAV 1171
Db 494 CLGMFPFCANPPGAVVGLLAGL-----VMAFW-----IGISL 536
QY 1172 VSAVGMVSR-----FVSHITRSPAIST---KPTWLERAKEATISMGSAVPA 1214
Db 527 VTMGSSMPPSPSNGSSPSLSLPTNLTVATVTTILMPLTTFKPTGLQR----- 579
QY 1215 GVAMTNLPGILVLGLAKAQLOIQIFFRNLNLTLLGLLHGLVPLPVLSVVGPDV 1274
Db 580 WYSANSTTVIVUGL-----IVSLLTGRMGRSLNPATITPVLPKLLSLLP 625
QY 1275 LR-QKR 1279
Db 626 LSCQKR 631

RESULT 75
US-11-096-568A-28283
; Sequence 28283, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28283
; LENGTH: 2186
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2186)
; OTHER INFORMATION: Ceres Seq. ID no. 2711348
US-11-096-568A-28283

Query Match 1.4%; Score 95.5; DB 7; Length 2186;
Best Local Similarity 19.7%; Pred. No. 1.3e+02;
Matches 189; Conservative 114; Mismatches 305; Indels 351; Gaps 48;

QY 570 LIMTSLNNYPAGDPRLL-----AQAKLWERAFLEEMRAFQRRMAGMQ 612
Db 478 LVQRFNSNGSYLDYIKLKLALETHERINTKVAASAASIWQSFSGATYTF----- 532
QY 613 VYTFABRSLEDEINRTTAEDLPFATSIVIVFLYISLALGSYSSSRVNVDSKATL 672
Db 533 SILKSSQSDQS-----TLDFVSVPKDIGTFTGLISGFLYTAWTMSKSR-GFG 579
QY 673 ---VAVVLGAVMAAMGFFSYLGISSILVQVVPVLVSVGADNIFIVLEVQRLP-- 725
Db 580 PMWVVFVGLVQVGFFFFTWASVVGLIAPPVPLMCL-----FVFLAGHS-LPFF 631
QY 726 -----RRPGBPREVHIGRALGRVAPS-MLLCSLSEAIC-----PFLGALT 765
Db 632 NVVTAARFSPQVGGTAVGIMQFGLSGAILQLTHAVCGGEGNPATFILLIAIV 691
QY 766 ---MPAVRTFALT-----GLA-----VILD 783
Db 692 FLAMPFVRVYETVITSDKKHDLGSAISMI IAAVLMVVITVENVLGSLSRM 751
QY 784 FLLQMSAFVALLSLSQKQESRLDVCCCKPQELPPPG-----QSGELLGLFFQ 838
Db 752 LLLASPLLVAVRALREKRTLSLSDGPVLDTSALLDPPSSNIFPPDGDHLV----- 805
```

```
QY 839 FLHWITRGVLLFLAL-----FGVSL-YSMCHISVGLDQELALPK-- 879
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 806 NILEAMSTVNFWLLFLAMLCMGSGFATVNNMRQIGESLRYs-----SVQLNSLSLWSIW 861
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 880 -----DSYLLDYFLFLNRYFEVGAPYVFTTILGYNFSSEAGNNAICSSAGCNFSFTQ 932
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 862 NFLGRFGAGYVSDTFLHKHSW---PRPIFMAITLGV-----MAIGHIIVASGVQGSLYAG 913
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 933 KIOVATEPPEQ-----SYLAIPASSWVDDFDLWLTSSCCRLYISG 973
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 SVLIGMAYGSQWSLMPITITSEIFGIRHMGTYIFTISIAGPIGSYI-----LSVKVIG 965
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 974 PNKDKFCPTVNSL---NCLKNCMSI-----TMGSVRPSVQF--HKYLPWFLNDRPNIK 1023
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 966 YFYDKVASEDDNCSFCGSCFRTSFMIMASVALFGSLVASVLFFRTHKFKYKNLVAKR-NLK 1024
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1024 CPKGLAAYS-----TSVNLTSDGQVLASRFMAYHKPLKN----- 1058
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1025 SLDSLSFAHABPPVDVQPHIVILSLVFPKVSAGVLKRDG--LALRLVNLVRLKSATPE 1082
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1059 -----SQDY-----TEALRAARELAANITAD-LRKVPGT 1086
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1083 CLISGLKCLVHLITTVESIMWNEGSDSYNILLNFVTHSDGKVRKCLASSCLRDVLQKSHGT 1142
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1087 DPAFEVFPYITNVFYEQVLTIL-----LPMLS-----LCLVPTFAVS 1127
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1143 -KAWQSVSGAITEMP-QNYLDLAHKSEVGSTEGARGAKQVLYILSTLKECLALMSKKHIA 1200
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1128 CILLGLDLRSGLNLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITR 1187
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1201 TLIEG-----FKVLMILRDPYITRPVID-SLNAVCL-NPTSEV--PVEALLEV-- 1244
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1188 SPAISTKPTWLERAKEATISMGSAVPAG-----VAMTNLPGILVLGLAKAQLIQIFPRL 1242
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1245 -----LSLAAGLFSGHETSADAMTFTARLLKVGWTRS-----FTL 1279
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1243 N---LLITLGLLHGLVFLPVLVSYPVDPNPALEOKRAEEAAV-----MVASC 1292
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1280 NRDLGVKLPSPVFNGL-----NDIIASEH---BEAIFAATDALKSLIFSC 1321
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ;
```

RESULT 76

```
US-11-096-568A-28282
; Sequence 28282, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28282
; LENGTH: 2199
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2199)
; OTHER INFORMATION: Ceres Seq. ID no. 2711347
US-11-096-568A-28282
```

Query Match 1.4%; Score 95.5; DB 7; Length 2199;
Best Local Similarity 19.7%; Pred. No. 1.3e+02;
Matches 189; Conservative 114; Mismatches 305; Indels 351; Gaps 48;

```
QY 570 LIMTFSLNYPAGDPRL-----AQAKLWEAFLEEMRAFORRMAGMFQ 612
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 491 LVQRFSNNSLYDYIKLKALETMERINTKWAAAASIIQSFSGATYTF-----AIYS 545
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 613 VTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSSYSWSRVVYDYSKATILGCG 672
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 77

```
US-11-096-568A-28281
; Sequence 28281, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28281
```

```
Db 546 SILKSSQSYDQS-----TLDFVSFXDIGCTFGIISGLFYLTAMTSKSR-GFGG 592
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 673 --VAVVLGAYMAAMGPPSYGIRSSSLVILQVVPFLVLSVGADNIFIFVLEYQRLP----- 725
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 593 PWVVVFGVLQVQFVGPFPIWASVVVGLIAPPVPLMCL-----FVFLAGHS-LPFFNTA 644
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 726 -----RRPGEPREVHIGRALGRVAPS-MLLCSLSSEAIC-----FFLGALTP----- 765
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 645 NVVTAARNFQYGGTAVGIMQGLGLSGAILIQIYHAVCGGEGNPATFILLAIIVPTLVM 704
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 766 --MPAVRTEALTS-----GLA-----VILD 783
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 705 FLAMPFVRVYETVITISDKHLDGLSALSMIIAAYLMVVITVENVLGLSRSMQIFSILVL 764
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 784 FLLQMSAFVALLSDSKRQENRLDVCCKVPQELPPG-----QGEGLLLGFFQKAYAP 838
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 765 LLLASPLLVAVRALREKRQTLSSLDGPVLDTSSALDPPSSNIFPDGDHLV-----AEDS 818
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 839 FLHWITRGVLLFLAL-----FGVSL-YSMCHISVGLDQELALPK-- 879
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 819 NILEAMSTVNFWLLFLAMLCMGSGFATVNNMRQIGESLRYs-----SVQLNSLSLWSIW 874
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 880 -----DSYLLDYFLFLNRYFEVGAPYVFTTILGYNFSSEAGNNAICSSAGCNFSFTQ 932
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 875 NFLGRFGAGYVSDTFLHKHSW---PRPIFMAITLGV-----MAIGHIIVASGVQGSLYAG 926
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 933 KIOVATEPPEQ-----SYLAIPASSWVDDFDLWLTSSCCRLYISG 973
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 927 SVLIGMAYGSQWSLMPITITSEIFGIRHMGTYIFTISIAGPIGSYI-----LSVKVIG 978
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 974 PNKDKFCPTVNSL---NCLKNCMSI-----TMGSVRPSVQF--HKYLPWFLNDRPNIK 1023
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 979 YFYDKVASEDDNCSFCGSCFRTSFMIMASVALFGSLVASVLFFRTHKFKYKNLVAKR-NLK 1037
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1024 CPKGLAAYS-----TSVNLTSDGQVLASRFMAYHKPLKN----- 1058
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1038 SLDSLSFAHABPPVDVQPHIVILSLVFPKVSAGVLKRDG--LALRLVNLVRLKSATPE 1095
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1059 -----SQDY-----TEALRAARELAANITAD-LRKVPGT 1086
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1096 CLISGLKCLVHLITTVESIMWNEGSDSYNILLNFVTHSDGKVRKCLASSCLRDVLQKSHGT 1155
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1087 DPAFEVFPYITNVFYEQVLTIL-----LPMLS-----LCLVPTFAVS 1127
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1156 -KAWQSVSGAITEMP-QNYLDLAHKSEVGSTEGARGAKQVLYILSTLKECLALMSKKHIA 1213
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1128 CILLGLDLRSGLNLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVEFVSHITR 1187
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1214 TLIEG-----FKVLMILRDPYITRPVID-SLNAVCL-NPTSEV--PVEALLEV-- 1257
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1188 SPAISTKPTWLERAKEATISMGSAVPAG-----VAMTNLPGILVLGLAKAQLIQIFPRL 1242
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1258 -----LSLAAGLFSGHETSADAMTFTARLLKVGWTRS-----FTL 1292
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1243 N---LLITLGLLHGLVFLPVLVSYPVDPNPALEOKRAEEAAV-----MVASC 1292
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1293 NRDLGVKLPSPVFNGL-----NDIIASEH---BEAIFAATDALKSLIFSC 1334
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 77

```
US-11-096-568A-28281
; Sequence 28281, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28281
```


; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: PCT/US2002/032263
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 10/287,994
; PRIOR FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-183-205-32

Query Match 1.4%; Score 95; DB 7; Length 461;
Best Local Similarity 21.0%; Pred. No. 16;
Matches 82; Conservative 38; Mismatches 126; Indels 144; Gaps 19;
QY 80 RLYTGPNQACCSAKQLVLSLASLITKALLTRCPACSD-----NFVN--LHCNHTCS 130
DB 43 REYDQTAQMCCS-KCSPGQHAKVCTKTSDTVCDSCEDSTYTQLMNWPCLSCGRCS 101
QY 131 PNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATAVLAVGTWCG 190
DB 102 SDQ-----VETQACTREQ-NRICTCRFGWYCA 127
QY 191 V---YGSALCNAQRWLNFGDGTGNGLAPLDTIFHLLPQAVGSGIQPLNEGVARCNSQ 247
DB 128 LSKQEGCRLC-----APL-----RKCRPG-----FGVAR-----P 152
QY 248 GDDVATCSQDCA-----ASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 153 GTETSDVVCPCAPGTFSTNTSTDICRPHQICNVV---AIPGN-----ASMDAVCTS 202
QY 301 LLVGRVA-----PARDKSMVDPKKGTSLSKLSF-----STHTLLGQF 340
DB 203 TSPTRSMAPGAVHLPOPVSTRSQHTQPTPEPSTAPSTFLLPMGFPSPPAEGST---GDF 258
QY 341 FQGWGTWASWPLTILVLSVIPVVALAAGLVTELTTPDV-----ELWSAPNSQARSEK 394
DB 259 ALPVLGVGTALGLLIIGVNCV-----IMTVKKKPLCLQREAKVPHLPADKARGTQ 312
QY 395 AFHQHGFPPFRTNQVILTAPNRSRYDS 424
DB 313 GPEQQH-----LLITAPSSSSSSLES 333

RESULT 80

US-11-260-192-2
; Sequence 2, Application US/11260192
; Publication No. US20060067934A1
; GENERAL INFORMATION:
; APPLICANT: SMITH, Craig A.
; APPLICANT: GOODWIN, Raymond G.
; APPLICANT: BECKMANN, Patricia M.
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR-ALPHA AND -BETA RECEPTORS
; FILE REFERENCE: A8466
; CURRENT APPLICATION NUMBER: US/11/260,192
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: US/10/420,785
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US 09/758,124
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 08/953,268
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: US 08/555,629
; PRIOR FILING DATE: 1995-11-09
; PRIOR APPLICATION NUMBER: US 08/468,453
; PRIOR FILING DATE: 1995-06-06

; PRIOR APPLICATION NUMBER: US 08/038,765
; PRIOR FILING DATE: 1993-03-13
; PRIOR APPLICATION NUMBER: US 07/523,635
; PRIOR FILING DATE: 1990-05-10
; PRIOR APPLICATION NUMBER: US 07/421,417
; PRIOR FILING DATE: 1989-10-13
; PRIOR APPLICATION NUMBER: US 07/405,370
; PRIOR FILING DATE: 1989-09-11
; PRIOR APPLICATION NUMBER: US 07/403,241
; PRIOR FILING DATE: 1989-09-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-260-192-2

Query Match 1.4%; Score 95; DB 7; Length 461;
Best Local Similarity 21.0%; Pred. No. 16;
Matches 82; Conservative 38; Mismatches 126; Indels 144; Gaps 19;
QY 80 RLYTGPNQACCSAKQLVLSLASLITKALLTRCPACSD-----NFVN--LHCNHTCS 130
DB 43 REYDQTAQMCCS-KCSPGQHAKVCTKTSDTVCDSCEDSTYTQLMNWPCLSCGRCS 101
QY 131 PNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATAVLAVGTWCG 190
DB 102 SDQ-----VETQACTREQ-NRICTCRFGWYCA 127
QY 191 V---YGSALCNAQRWLNFGDGTGNGLAPLDTIFHLLPQAVGSGIQPLNEGVARCNSQ 247
DB 128 LSKQEGCRLC-----APL-----RKCRPG-----FGVAR-----P 152
QY 248 GDDVATCSQDCA-----ASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 153 GTETSDVVCPCAPGTFSTNTSTDICRPHQICNVV---AIPGN-----ASMDAVCTS 202
QY 301 LLVGRVA-----PARDKSMVDPKKGTSLSKLSF-----STHTLLGQF 340
DB 203 TSPTRSMAPGAVHLPOPVSTRSQHTQPTPEPSTAPSTFLLPMGFPSPPAEGST---GDF 258
QY 341 FQGWGTWASWPLTILVLSVIPVVALAAGLVTELTTPDV-----ELWSAPNSQARSEK 394
DB 259 ALPVLGVGTALGLLIIGVNCV-----IMTVKKKPLCLQREAKVPHLPADKARGTQ 312
QY 395 AFHQHGFPPFRTNQVILTAPNRSRYDS 424
DB 313 GPEQQH-----LLITAPSSSSSSLES 333

RESULT 81

US-11-087-099-11321
; Sequence 11321, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11321
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-11-087-099-11321

Query Match 1.4%; Score 95; DB 7; Length 469;
Best Local Similarity 25.9%; Pred. No. 16;
Matches 42; Conservative 19; Mismatches 51; Indels 50; Gaps 7;

QY 226 GOAVGSGIQPLNEGVARCNSQGDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSL 285

Db 18 GNMVSGIFSLPSSLA-----SIASPFQATSAWLL-----TGAG 51
Qy 286 VLIILICSVFAVVTILLNGFRVAP-ARDKSKWDPKGTSLSKLSFSHTLLGQFFQW 344
Db 52 VLMIAL--VFHLSIRKPELTACQSYARALFSDPKKNAAG-----FTWVW 96
Qy 345 GTWVASWPTIIVLSVIPVVALAAGLV--FTBLTDDPVELMS 384
Db 97 GYVVASW-----ISNVAITSLAGYLTSFFPILVDKREMF 132
RESULT 82
US-11-087-099-10983
; Sequence 10983, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; PRIOR FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10983
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Methanosarcina barkeri
US-11-087-099-10983

Query Match 1.4%; Score 95; DB 7; Length 552;
Best Local Similarity 17.4%; Pred. No. 20;
Matches 109; Conservative 82; Mismatches 185; Indels 250; Gaps 30;
Qy 635 IPATSVIVIFLYIS-----LALGSYSWSRWVDSKATLGLG-----GVAV 675
Db 52 VAGTMYIISLCASLGSKATLAGPY-DWAKLFTGPGAAVGLAVMEYALREADAI 110
Qy 676 VLGA-----VMAAMGFFSYLGIRSLV---ILQVPELVLSVGADNIF 715
Db 111 VVASISQSIFPELQVVPVTLVIALTFINRGVVAALTLNVLWTIAFIALA---F 165
Qy 716 IFVLYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEAICFFIGALTTPMPAVRTFALT 775
Db 166 PFSTAF-----GIGDIHPEYLL-----QCAL-PNGMIGLEA---195
Qy 776 SGLAVILDPLQMSAFVALLSDSKRQEARLDVCCVKPQELPPGQEGELGFFQKA 835
Db 196 -----ALQFGPWF-YLGIEGAMCAE-----CKHPSRAVPLGQOAGMIT---234
Qy 836 YAPFLLHMTIRGVVLLFLA-----LFGVLSYSMCHISVGLDQELALPKDSYLLDYFL 888
Db 235 -----LLIGAAMTLYLCSVLIPADLLGVSVYPL-----262
Qy 889 FLNRYFEV---GAPVYVVTLLGN---FSSEAGMN-AICSSAG---CNNPSFTQKIYATE 939
Db 263 -----FEAQNQGVFIYVALLGLGTFLTCVASANGCVCDSSRLVCSLKRQLRVIRFSAV 317
Qy 940 PPEQS--YLAI-----PASSWDDDFIDWLTPSSCCRLYISGPNKDKFCPTVNSL 987
Db 318 HPKNTPTNAVIFTPVPAIGFAFGDLQDVIIF-----SIVSGL 356
Qy 988 NC-----LKNCMSITMGSVRPSVEQPHKLPFWFLNDRPNIKCPKGGIAAYSTSVN 1037
Db 357 LCYVLIPFLIRFKLFPETTSKVRFPVGLQPIYAF-----AIAIA 399
Qy 1038 LTSQGVILASRFMAYHKPL-----KNSQDYTEALRA-----ARELA 1073
Db 400 IT-----ILSTLFWGKYKYLIFAFVSVLRTSISATNTRQTLKIGLKWAGPYPKAREIE 455
Qy 1074 ANITADLRKVPCTDPAFEPPTIYNVFEQYL--TILPEGLFMLSCLVPTFAVSCILL 1131
Db 456 RSRRLQ-----ALETNDQALSKYHSLKTGDTALVVGMLALLICVGEFIAYKILSE 507

Qy 1132 GUDLRSGLLNLSI---VMILVDTVG 1154
Db 508 TWEIASGPFYLYVIPTGLIVLIETIG 533
RESULT 83
US-10-505-263-8
; Sequence 8, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; PRIOR FILING DATE: 2004-08-20
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-505-263-8

Query Match 1.4%; Score 95; DB 6; Length 735;
Best Local Similarity 20.1%; Pred. No. 31;
Matches 124; Conservative 72; Mismatches 202; Indels 220; Gaps 32;
Qy 719 LEYQRLPRRPGSPREKHIGRA-----LGRVAP-----SMLCSLSEAICFF-- 759
Db 1 MELQR-----RDYHVERPLLNQLEDLGHGWPAAKTHQWRTWFRCSRARHSLLLQ 52
Qy 760 -----LGALTTPMPAVTFA---LTSGLAVILDPLQMSAFVALLSDSKRQEARLDVCC 812
Db 53 HVPVLGLPRYP-VREWLLGDLGLSLVAIMQLPQGLAYALLAGL-----96
Qy 813 VKPQELPPGQEGELGFFQKAYAPFLLHMTIRGVVLLFLALFGVLSYSMCHISVG-- 870
Db 97 -----PP-----MFLYSSFPYFIYP-----LFTSR-----HISVGTF 126
Qy 871 -----LQDELALPKD-----SYLLDYFLNRYFEVGPVY---P 902
Db 127 AVMSVMVGSVTESLTADKAFVQGLNATADDDARVQVAYTLLSLVGL---FQVGLGVHFGF 183
Qy 903 VTT-----LGVNFSSEAGMNAICSSAGCNPSFTQKIYATEPPEQSYLAIPASSWDDP 957
Db 184 VVTVLSBFLVRSYTTAASVQVLVS-----QLKYVFGIKLSSHSG-PLS-----225
Qy 958 IDWLTPSSCCRLYISGPNKDKFCPTVNSLKNCMSITMGSVRPSVEQPHKLPFWFLN 1017
Db 226 VIYTVLEVCAQL-----PETVGTVVTAIVAGVALVVLKLLNEKLRRLP---- 270
Qy 1018 DRPNIKCPKGGIAAYSTSVNLTSDQGVILASRFMAYHKPKNSQDYTEALRAARELANIT 1077
Db 271 -----LPIGBELLTIGATGISYGVKLNDRPKV-----DVVGNIT 305
Qy 1078 ADLRKVPCTDPAFEPPTIYNVFEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLRS 1137
Db 306 TGL--IPPVAPKTELFAVLGNAP-----AIAVVG-PAIAISLGKIFA---LRHGYRVS 354
Qy 1138 -----GLNLSLIVMILVDTVGFMALWDISYN-AVSLINLVASVGMSEVFSHTTRSF 1189
Db 355 NOELVALGLNLDIG-----GPFQCFPVSMSRSLVQ--ESTGGTQVAGAVSSLP 403
Qy 1190 AISTKPTWLERAKEATISMGSAVFAGVAMTNLPGL-----VLGLAKAQLIQIFPRLNL 1244

Db 483 LSLALISIMGLICCCQOKT---GLPWAFLAVLLAWLITTFACWYGLVGY----- 533
QY 653 SYSSWSRVMVDSKATLGLGAVVAVLAVM-----AAMGFFSYLGRSSVLVQVVPFLVL 707
Db 534 -YOPTSAVO-----MIGAYMVRPVPVMMFTLYGSAVQAIQMLGDLKL 578
QY 708 SVGADNIFVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEALCFGLALTPMP 767
Db 579 A-----QYAKLP-----PRATFLAQILGTGAGSVFNMVWNSI----- 611
QY 768 AVRTFALTGLAVILDFLLQMSAFVALLSLDSKQESRLDVCCVK----- 814
Db 612 -----VDNR-----DILLSVQGNVWSGQNVQY 636
QY 815 -PQELPPPGQEGLL-----LGFFOKAYAPFLHMITRGVVLVLLFLALFGVS 860
Db 637 NSQAVAWGGTANEILVHSGRHGTYVMVPMGLFGFIAP-LPWI--GHKYFPQLRLNHIN 693
QY 861 LYSMC-----HISVGLDQELALPKDSYLLDYFL-----LNRYFVGAPVVPVTTLGYN 909
Db 694 TFLICTWGLSVGI-----NSSLLAYFVGFPSQGLRY-----KPALE----- 734
QY 910 PSSEAGMNAICSSA---GCN-----NFSFTQKIYATEPQSYLAIPASSW-----VDD 956
Db 735 -----AKWNNICAAIAGGCSLIIFILTFVAGSGKARDPPK-----WNGNMDG 780
QY 957 FIDMLTPSSCCRLYISG 973
Db 781 NVD-----RCKMYLNG 792

RESULT 86

US-11-087-099-4589
; Sequence 4589, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4589
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Bacteroides thetaiotaomicron VPI-5482
US-11-087-099-4589

Query Match 1.4%; Score 94.5; DB 7; Length 231;
Best Local Similarity 21.9%; Pred. No. 6.4;
Matches 42; Conservative 35; Mismatches 70; Indels 45; Gaps 6;
QY 1106 LILPGLFMLSCLVPTTAVSCLLGLDLRGLNLSIVMLVDVTGFMALMDSYNA 1165
Db 1 MSFLENPFLLAV---TEGVFFAKLQKTKGLVLLNPILLITIAVLIIFLKMANISYET 56
QY 1166 VSLINLVSAVGSVERPVSHITRSPASTKPTWLERAKEATISNGSAVFAGVATNLPGL 1225
Db 57 -----YNOGHHLIEF-----WL---RPAVVALG-----VPLYL 81
QY 1226 VLGLAKAQILQIIFFRFLNLLITLGLHGLVFLPVLISYVGVDPVNPALAEOKRAEEAVA 1285
Db 82 QLEMICKQLLP-----LLSQLAGCIVGVISVLIKFPWASQEVLSLAPKSVTTPIA 135
QY 1286 AVWVASCPNHS 1297
Db 136 MEVTKAIGGIPS 147

RESULT 87

US-11-087-099-2054
; Sequence 2054, Application US/11087099

; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2054
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Methanosarcina barkeri
US-11-087-099-2054

Query Match 1.4%; Score 94.5; DB 7; Length 467;
Best Local Similarity 19.7%; Pred. No. 18;
Matches 78; Conservative 53; Mismatches 140; Indels 125; Gaps 19;
QY 635 IPATSYIVITFLYIS-----LALGSYSSWSRVMVDSKATLGLG-----GVAV 675
Db 52 VAGTMYVIIISLCASBLGSAKLGGPY-DWARLFIGPAAASVGLAVYMEYIALEADA 110
QY 676 VLGA-----VMAAMGPFESYLGRSSLV---ILQVVPFLVLSVGADNIF 715
Db 111 VVASISQSIPELQVVPVTLVIALLTFINYRGVAAALTINFLVIAFIALLA-----F 165
QY 716 IPVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEALCFGLALTPMPAVRTFALT 775
Db 166 FESTA-----GTDIHPYLL-----QGLL-PNGMIGLFA-- 195
QY 776 SGLAVILDFLQMSAFVALLSLDSKQESRLDVCCVQBELPPPGQEG-----LLGPF 832
Db 196 -----ALQFGPWF-YLIGEGAMCABE-----CKHPSRAVPLGQAGMITLLIGAA 240
QY 833 QKAY-----APFLHWTIRGVVLLFLALFGVLSYMSCHISVG-----LDQE-- 874
Db 241 MTLYLCSVLIPADLLGVSVYPLFEAAQNSGIFIFVALLGLGTFLTCVASANGCVCDSRS 300
QY 875 -LALPKDSYLLDYFLPLN-RYPEVGAPVYFVTTLGYNPFSEAGMN-----AICSSAGCNN 927
Db 301 WFLASRDNVVSSWFSANVHPKYNTPRAVITVPVAGFAPSGVLDQVITFSIVGLLCYV 360
QY 928 FSFTQKIYATEPQSYLAIPASSWVDDFIDWLT 963
Db 361 LIPFSLIRFKLFE-----TSKVRPFPVGLQP 389

RESULT 88

US-10-479-873-9
; Sequence 9, Application US/10479873
; Publication No. US20060024673A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC2A5 AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-076
; CURRENT APPLICATION NUMBER: US/10/479,873
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-479-873-9

Query Match 1.4%; Score 94.5; DB 6; Length 477;
Best Local Similarity 22.5%; Pred. No. 18;

Matches 108; Conservative 54; Mismatches 157; Indels 161; Gaps 27;
QY 526 ANAPLTFKDGKALASCMADYCAPVPFLAIGGYGKDYSEAEALIMTFSLNYPAGDPR 585
Db 34 ALGPLSF--GPAFG--YSSPAIP-----SLQRAAPAPR 63
QY 586 LAQAKL--WEEAFLEBRAPQORMAGFQVTTAERSLEDEINRTTAEIDLPIFATSYIVIF 644
Db 64 LODAASWFGAVVTGAAAGGVIGW-----LVDRAGRLS---LLCSVPFVAGP 111
QY 645 LYISIALGYSYSSWRSVMVDSKATLGLG--GVAVVLGAV--MAAMGFYSYLGIRSSLVILQV 702
Db 112 AVITAA---QDWW--MLGGRLLTGLACVASLAPVVISEIAYPAVRGLGSCVOLMV 166
QY 703 PFLVLSGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGA 762
Db 167 VGILLAYLAG--WVLEWRML-----AVLGCVPPLMLLL---LMCF----- 201
QY 763 LTPMPAVRTEALTSLAVILDFLLQMSAFVALLSLDSKQEBASRLDVCCVKPQELPPPG 822
Db 202 ---MPETPRFLUTQH-----RQEAAMALRFLWSEQ-----GWEDDPFG 238
QY 823 QGEGLLGFFQK--AYAPFLHWTIRGVVLLIFLALFGVSLY-----SM 864
Db 239 AEQSHLALLRQPGIYKPII-----GVSLMAFQQLSGVNAVMFYAETIFBEAKPKDSSL 293
QY 865 CHISVGLQDELAKPKDSYLLD-----YFLNRYFEVGAAPVYVTT--LQYNFSSBAGNA 918
Db 294 ASVWGVQIVLPTAVALIMDRAGRLLLVLS-----GVVMVFSTSAFGAYFKLTQG--- 345
QY 919 ICSSAGCNFTSQIKIYATEPEOSYLAI--PASSWDDF---IDWLTPSSCCRLYISG 973
Db 346 -----GPGN-----SSHVAISAPVSAQPDVDSVGLAWLAVGSMC--LFIAG 384

RESULT 89

US-11-087-099-5936
; Sequence 5936, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:

; APPLICANT: Abad, Mark S., et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464

; SEQ ID NO 5936

; LENGTH: 599

; TYPE: PRT

; ORGANISM: Ferroplasma acidarmanus

US-11-087-099-5936

Query Match 1.4%; Score 94.5; DB 7; Length 599;

Best Local Similarity 22.2%; Pred. No. 25;

Matches 125; Conservative 66; Mismatches 171; Indels 201; Gaps 33;

QY 283 GSVLIIILCSFPAVVTILLGVFPAPADKSKMDPKGTSLSDKLSSTHTLLQGFQ 342

Db 88 GAIVLVVL--VYAEICAMI-----PRSG--FITRYGHYSHGGIAGLPF 127

QY 343 GNGTWWASPLTILVLSVTPVVALAAGLVFT--ELTTDPVELMSAPNSQARSEKA----- 395

Db 128 GWGYFAAR-----VAAPALAEAAITAGSYITKPAIYYAKNPPLDPSSSVTLISY 179

QY 396 -----PHDQHFGP--PFRTNQVI-----LTAPNRSYRYOSLLLGPKNFS----- 433

Db 180 GILIAAALTAGFFLYNFGVLMGKTNQGITWKKLIIPSTIILMVFLLPHAGNFNNPAL 239

QY 434 -GIL-----DLDLLELLELQ-----RLRHLQWSPEA---QRNISLDICVAPLNPDT 480

Db 240 GGFLPHNNISLVFEAISTDGVFSYLGFRHTLNFAGEAKNPQDIP--RALIYAML-----T 294

QY 481 SLYDCCINSILQY--FQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTPKOGTALA 539

Db 295 SI---VVVLLQAFISAINPSLLTASGWLGLSS-----ASA 329
QY 540 LSCMADYCAPVPFLAIGGYGKDYSEAEALIMTFSL--NNY--PAGDPLAQAALWEEA 595
Db 330 GTYAKSIDSAFFAPLA-----KSSSLAILAVLYLLYDAYTSPAGTLNIAAGTATRSL 383
QY 596 F-LEEMRAFQORMAGFQVTTAERSLEDEINRTTAEIDLPIFATSYIVIFLYISIALG-- 652
Db 384 YGLAEI-----GYFPRIF-----GKSKRTG--VPVFS-----LLISLVGLI 419
QY 653 ---SYSSWSRVMVDSKATLGLGVAVVLGAVMAAMGFYSYLGIRSSLVILQVPLVLSV 709
Db 420 FLVLPSPW-----YVVVGLVSGVAGFYILG--GSTLMVLR----- 453
QY 710 GADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGA-----LT 764
Db 454 -----REASDLKR-----PFKLPYARILSPVA-----FIGASLIVVTW 486
QY 765 EMPAVRTPA--LTSGLAVILDFL 785
Db 487 GWPTVAYIAIILFAGFAVLVFL 509

RESULT 90

US-11-205-109-4

; Sequence 4, Application US/11205109

; Publication No. US20050287641A1

; GENERAL INFORMATION:

; APPLICANT: Farnet, Chris

; APPLICANT: Zazopoulos, Emmanuel

; APPLICANT: Staffa, Alfredo

; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS

; FILE REFERENCE: 3002-ZUS

; CURRENT APPLICATION NUMBER: US/11/205,109

; CURRENT FILING DATE: 2005-08-17

; PRIOR APPLICATION NUMBER: US/09/976,059

; PRIOR FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: US 60/239,924

; PRIOR FILING DATE: 2000-10-13

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 4

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Actinoplanes sp.

US-11-205-109-4

Query Match 1.4%; Score 94; DB 7; Length 336;

Best Local Similarity 23.1%; Pred. No. 12;

Matches 68; Conservative 44; Mismatches 126; Indels 56; Gaps 13;

QY 596 FLEEMRAFQORMAGFQVTTAERSLEDEINRTTAEIDLPIFATSYIVIFLYISIALG----- 651

Db 65 YLAGVALVPGLLGCGFWGAPLITRELENGTQRLVWVNSVTRRRRLWLLKLLVVGLACMVVA 124

QY 652 ---GSYSWSRVMVDSKATLGLGVAVVLGA---VMAAMGFYSYLGIRSSLVILQVVPF 704

Db 125 GVPSLLLTWAAAPVDNVADNRF--TVMFGARFLPIAYAAFAVLGTLLGLLVRETVA 182

QY 705 LVLSVGADNIFIVLEYQRLPR--RPGPREVHIGRALGRVAPSMMLCSLSEACFFLGA 762

Db 183 MALTLVAVFVIFQFL-----VPLNVRP-----HLPAPKHLVXP-MTVSAINEAKS--LGS 228

QY 763 LTPMPAVRTEALTSLAVILDFLLQMSAFVALLSLDSKQEBASRLDVCCVKPQELPPPG 822

Db 229 ITGAPVL-----NGLSISQGWITDVS---ALKTADGRSLDAKTFDNCY-----MNAKTKG 275

QY 823 QCEG-----LLLGPFQKAYAPFLHWM-----ITRGVWLLLLFLALFGVSLY 862

Db 276 ATEPGYGDVAVCLAKLDLHDVIAYQPNRYWAFQFLESFYVLLSGLLGAAMV 329

[illegible]

```

RESULT 94
US-11-087-099-2925
; Sequence 2925, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2925
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Shigella flexneri 2a str. 2457T
US-11-087-099-2925

Query Match      1.4%; Score 93.5; DB 7; Length 461;
Best Local Similarity 18.6%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 159; Indels 227; Gaps 24;

Qy 752 LSAEACFFGALTPMPAVENTALTSTGLA---VILDPELLOMSA--FVALLSLDSKREASR 806
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 LQQVMWGLAYLTPMTVDFTEGIVSGISGHWPASTYLLALAGLVFAISYGKLVRO---- 79

Qy 807 LDVCCCVKQELPPPGQEGELLGFFQKAYAP---FLLHWITRGVVLLFLALFGVSLYS 863
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 -----PPEAGSA---YTYAQKSNINPHVGPWGM----- 104

Qy 864 MCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAQVYFTVILGYNFSSEAGMNAICSSA 923
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 -----SSLLDY-LFL-----PM-----INVLLA-- 121

Qy 924 GCNFSFTQKIYATPEPQSYLAIPASSWVDDFDWLTPSSCCRLYISGPNKDKFCPST 983
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 -----KIYLSALFPE-----VPPWVWVTFVAILTAAN----- 149

Qy 984 VNSLNCILKNC-----MSITGMSVRPSVEQPHK-----YLPWFNLDRPNIKCPKG 1027
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 LKSNVLVANFTLFLVLQVISVFLVQVGLHKGEGVGTVMWSLOQPFISENAHL----- 204

Qy 1028 GLAAYSTVNLTSQGQVLASRFWAYHKPLKNSODYTEALRAARELAANITA----- 1078
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 -----IPITGATIVCSFSGFDNAVTLTSEETPDAARVIPK-AIFLTAVYGVGFIA 255

Qy 1079 -----DLARKVPGTDPAFVFPFYITINVFYEQYLTLPE-----GLFMLSCLVLP 1122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 256 ASPEFMQLFFPDISRFKDPDAA-----LPEIALYVGGKLFQSFPLCT 296

Qy 1123 TFVAVSCLLGLDRLSGLLNLSIV-----MTLVDVTGFMAL 1158
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 TF-VNTLASGLASHASRLLYVNGRDNVPPERFVGVHPKWTPTALNVTMVGIVALSAL 355

Qy 1159 WDISYNAVSLIN---LYSAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAG 1215
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 FFDLVLTATALNFGALVAFTFVNLSVFNHFWRRKGMN--KSWKDHFYLLMPLVGLTVG 413

Qy 1216 VAMTNLPGEI-LVIGLAKAQL 1234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 VLWVNLESTSLTLGLVWASL 433

RESULT 95
US-11-087-099-4376
; Sequence 4376, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464

```

```

; SEQ ID NO 4376
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Es
US-11-087-099-43

```

Query Match 1.4%; Score 93.5; DB 7; Length 461;
Best Local Similarity 18.6%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 159; Indels 227; Gaps 24;

[illegible]

RESULT 96

```

US-11-087-099-9669
; Sequence 9669, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad Mark S. et al.
; TITLE OF INVENTION: Genes and Uses fo
; FILE REFERENCE: 38-21(53450) B EP
; CURRENT APPLICATION NUMBER: US/11/087
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9669
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7 E
US-11-087-099-9669

```

Query Match 1.4%; Score 93.5; DB 7; Length 461;
Best Local Similarity 18.6%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 159; Indels 227; Gaps 24;
QY 752 LSEAICFFGLTPMPAVRTALTSLA----VILDFLLMSA--FVALLSDSKROBSR 806

Db	24	LMQVVMGLAYLTMTVFTDFGIVSGISDGHVPASVYLLALAGVLTFAISYKGLVRQ	79
Qy	807	LDVCCCKPQBLPPGQEGLLGFFKQAVP---FLLLHWTTRGVLLLLFLALFGVSLYS	863
Db	80	-----FPEAGSA---YTYAQKSNPHVGFVMG	104
Qy	864	MCHISVGLDQELALPKDSYLLDYPFLNRYFEVCGAPVFTVLGYNFSSEAGMNAICSSA	923
Db	105	-----SSLLDY-LFL-----PM-----INVLLA--	121
Qy	924	GCNNFSFTQKIQYATEPPEOSYLAIPASSWVDDFDMLTPSSCCRLYISGPNKDKFCPST	983
Db	122	-----KIYLSALPPE-----VPPWVWVTFVAILTAAN-----	149
Qy	984	VNSLNCUKNC-----MSITGWSVRPSVEQFHK-----YLPWFLNDRPNIKCPKG	1027
Db	150	LKSNVLVANFTLFLVLQISIMVFIPLAVQGLHKGEGVGVWLSQPFISENAHL-----	204
Qy	1028	GLAAYSIVNLTSDGQVLSRFPWAYHKPLKNSQDYTEALRAARELANITA-----	1078
Db	205	-----IPITGANTIVCFSEFLGDFAVTTLSEBTPDAARVIPK-AIFLTAVYGVGFIA	255
Qy	1079	-----DLRKVCTDPAFEPVPPYITNIFYBQVLTILPE-----GLFMLSCLIVP	1122
Db	256	ASPFMQLFPDISRFKDPDAA-----LPEALYVGGKLFQSFIFLCT	296
Qy	1123	TFAVSCILLGLDRSGLLNLSIV-----MILVDTVGFMAL	1158
Db	297	TP-VNTLASGLASHASVSRLLYVGRDNVFPERVGVHPKWRTPALNVIMVGLIVMSAL	355
Qy	1159	WDISYNAVSLIN---LVSAYGMSVEFVSHITRFSFAISKTPTWLERAKEATISMGSAVPAG	1215
Db	356	FEDLVATATALINFGALVAFTFVNLSVFNHFWRRKGMN--KSKDKHFFHYLLMPLVGLATVG	413
Qy	1216	VAMTNIPLGI-LVJGLAKAQL	1234
Db	414	VLWVNLESTLTGLVWASL	433

RESULT 97

```

US-11-087-099-10340
; Sequence 10340, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Pl
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10340
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(602)
; OTHER INFORMATION: unsure at all Xaa loc
US-11-087-099-10340

```

Query Match 1.4%; Score 93.5; DB 7; Length 602;
Best Local Similarity 23.8%; Pred. No. 30;
Matches 60; Conservative 36; Mismatches 87; Indels 69; Gaps 13;

[illegible]

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 4868
; TYPE: PRT
; ORGANISM: Meleagris gallopavo
US-11-044-111-24

Query Match 1.4%; Score 93.5; DB 7; Length 4868;
Best Local Similarity 20.2%; Pred. No. 5.9e+02;
Matches 80; Conservative 63; Mismatches 121; Indels 133; Gaps 25;

QY	893	YPEVGAPVYFTTLGYNFSSEAGMNAICSSAGCANNFSTQKIQYATEPPEQS-----YLA	947
Db	1286	YCLSLNPTEFRSS--FNFG--VGVENASSDA-----LQKRKHSOEFPASSTTYFYSLR	1334
QY	948	I-----PASSWDDFDIMLTPSSCCRLYISGPNKDKFCPCSTVNSLNCNK-MSITMGS-	1000
Db	1335	IPAGQDPSSVWV---GWVTPD---YHFYSENF-----IN-----KNCIVTVTLGDE	1375
QY	1001	---VRPSVEQFHKLWP-----FLNDRPNIKCPKGGLAAYSTS-VNLTSQGQVLASR	1048
Db	1376	RGKVHESVRSNCYMWGCGDITANSQSRGSRNVDLFEGCFVDLATGMLSTANGKELGTC	1435
QY	1049	EM-----AYHKP-----LKNQSDYTEALRAARELAANITADLRKVPCT	1086
Db	1436	YQVEPNTKLLPAFVQPTSTNLIQFELGKIKNTMPLSAAIKSEE-----RNPVQC	1487
QY	1087	DPAFEVFPYTIITNVFEQYLITLPEGLFMLS-----CLVPTFAVSCLLILGLD	1134
Db	1488	PPLDVL-QTIIPVLSR---MPNSFLKVEVERSERHGWVQCLEP---LQMMALHIP	1538
QY	1135	LRSGLNLNLSIVMILVDTVGFMAWDISYNAVSLINLVSAGV---MSVEFVSHITRS---	1188
Db	1539	EENRCVDILE---LCEQEDLM---KPHYHTLKLYSSVCALGNTRVAYALCSHVDISQLF	1591
QY	1189	FAITK-----PTWLERAKEATISMG	1210
Db	1592	YTIQNYLPGLLRSGFYDLLISIHLDHAKQAKLMNN	1628

Search completed: April 11, 2006, 01:01:32
Job time : 53 secs

THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:46:22 ; Search time 556 Seconds
(without alignments)
3310.727 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAIISFLPNNGROF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US066 COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US073 COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US074 COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US075 COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US076 COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US077 COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US078 COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US079 COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US080 COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US097 COMB.pcp.*
28: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
29: /cgn2_6/ptodata/1/paa/US099 COMB.pcp.*
30: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
31: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
32: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
33: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
34: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
35: /cgn2_6/ptodata/1/paa/US105 COMB.pcp.*
36: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
37: /cgn2_6/ptodata/1/paa/US107 COMB.pcp.*
38: /cgn2_6/ptodata/1/paa/US108 COMB.pcp.*
39: /cgn2_6/ptodata/1/paa/US109 COMB.pcp.*
40: /cgn2_6/ptodata/1/paa/US110 COMB.pcp.*
41: /cgn2_6/ptodata/1/paa/US111 COMB.pcp.*
42: /cgn2_6/ptodata/1/paa/US112 COMB.pcp.*
43: /cgn2_6/ptodata/1/paa/US114 COMB.pcp.*

44: /cgn2_6/ptodata/1/paa/US600 COMB.pcp.*
45: /cgn2_6/ptodata/1/paa/US601 COMB.pcp.*
46: /cgn2_6/ptodata/1/paa/US602 COMB.pcp.*
47: /cgn2_6/ptodata/1/paa/US603 COMB.pcp.*
48: /cgn2_6/ptodata/1/paa/US604 COMB.pcp.*
49: /cgn2_6/ptodata/1/paa/US605 COMB.pcp.*
50: /cgn2_6/ptodata/1/paa/US606 COMB.pcp.*
51: /cgn2_6/ptodata/1/paa/US607 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	1 PCT-US03-22467-4	Sequence 4, Appli
2	6909	100.0	1332	1 PCT-US03-40113-4	Sequence 4, Appli
3	6909	100.0	1332	1 PCT-US05-01469-4	Sequence 4, Appli
4	6909	100.0	1332	36 US-10-621-758A-4	Sequence 4, Appli
5	6909	100.0	1332	36 US-10-646-301A-4	Sequence 4, Appli
6	6909	100.0	1332	36 US-10-663-208A-4	Sequence 4, Appli
7	6909	100.0	1332	37 US-10-736-769-4	Sequence 4, Appli
8	6909	100.0	1332	37 US-10-750-386-4	Sequence 4, Appli
9	6896	99.8	1332	1 PCT-US01-04098A-1831	Sequence 1831, Ap
10	6896	99.8	1332	32 US-10-239-316-9	Sequence 9, Appli
11	6896	99.8	1332	32 US-10-258-899A-1831	Sequence 1831, Ap
12	6896	99.8	1332	32 US-10-293-244-1831	Sequence 1831, Ap
13	6896	99.8	1332	42 US-11-218-141-1831	Sequence 1831, Ap
14	6896	99.8	1332	42 US-11-242-459-9	Sequence 9, Appli
15	6872.5	99.5	1359	1 PCT-US01-04098A-1830	Sequence 1830, Ap
16	6872.5	99.5	1359	1 PCT-US03-22467-44	Sequence 44, Appli
17	6872.5	99.5	1359	1 PCT-US03-40113-44	Sequence 44, Appli
18	6872.5	99.5	1359	1 PCT-US05-01469-44	Sequence 44, Appli
19	6872.5	99.5	1359	1 PCT-US05-27579-22	Sequence 22, Appli
20	6872.5	99.5	1359	32 US-10-258-899A-1830	Sequence 1830, Ap
21	6872.5	99.5	1359	32 US-10-293-244-1830	Sequence 1830, Ap
22	6872.5	99.5	1359	36 US-10-621-758A-44	Sequence 44, Appli
23	6872.5	99.5	1359	36 US-10-646-301A-44	Sequence 44, Appli
24	6872.5	99.5	1359	36 US-10-663-208A-44	Sequence 44, Appli
25	6872.5	99.5	1359	37 US-10-736-769-44	Sequence 44, Appli
26	6872.5	99.5	1359	37 US-10-750-386-44	Sequence 44, Appli
27	6868.5	99.4	1359	31 US-11-218-141-1830	Sequence 1830, Ap
28	6868.5	99.4	1359	48 US-60-453-050-8956	Sequence 8956, A
29	6868.5	99.4	1359	48 US-60-466-412-8956	Sequence 8956, A
30	6868.5	99.4	1359	48 US-60-466-412-8956	Sequence 8956, A
31	6868.5	99.4	1359	48 US-60-466-412-8956	Sequence 8956, A
32	6536	94.6	1344	1 PCT-US01-08631-53052	Sequence 53052, A
33	6536	94.6	1344	34 US-10-450-763-53052	Sequence 53052, A
34	5421.5	78.5	1331	1 PCT-US03-22467-2	Sequence 2, Appli
35	5421.5	78.5	1331	1 PCT-US03-40113-2	Sequence 2, Appli
36	5421.5	78.5	1331	1 PCT-US05-01469-2	Sequence 2, Appli
37	5421.5	78.5	1331	36 US-10-621-758A-2	Sequence 2, Appli
38	5421.5	78.5	1331	36 US-10-646-301A-2	Sequence 2, Appli
39	5421.5	78.5	1331	36 US-10-663-208A-2	Sequence 2, Appli
40	5421.5	78.5	1331	37 US-10-736-769-2	Sequence 2, Appli
41	5421.5	78.5	1331	37 US-10-750-386-2	Sequence 2, Appli
42	5407	78.3	1333	1 PCT-US03-22467-12	Sequence 12, Appli
43	5407	78.3	1333	1 PCT-US03-40113-12	Sequence 12, Appli
44	5407	78.3	1333	1 PCT-US05-01469-12	Sequence 12, Appli
45	5407	78.3	1333	1 PCT-US05-27579-3	Sequence 3, Appli
46	5407	78.3	1333	36 US-10-621-758A-12	Sequence 12, Appli
47	5407	78.3	1333	36 US-10-646-301A-12	Sequence 12, Appli
48	5407	78.3	1333	36 US-10-663-208A-12	Sequence 12, Appli
49	5407	78.3	1333	37 US-10-736-769-12	Sequence 12, Appli
50	5407	78.3	1333	37 US-10-750-386-12	Sequence 12, Appli
51	4466	64.6	982	1 PCT-US01-08631-53050	Sequence 53050, A
52	4466	64.6	982	34 US-10-450-763-53050	Sequence 53050, A
53	2403	34.8	1318	39 US-10-940-774-10152	Sequence 10152, A
54	2403	34.8	1318	39 US-10-940-774A-10152	Sequence 10152, A
55	2402.5	34.8	1278	1 PCT-US03-40977-530	Sequence 530, App

56 2402.5 34.8 1278 1 PCT-US03-40978-1542 Sequence 1542, Ap
57 2402.5 34.8 1278 1 PCT-US04-07141-810 Sequence 810, App
58 2402.5 34.8 1278 1 PCT-US05-18850-1032 Sequence 1032, App
59 2402.5 34.8 1278 31 PCT-US05-18850-1032 Sequence 35088, A
60 2402.5 34.8 1278 32 US-10-170-205E-35088 Sequence 2, Appl1
61 2402.5 34.8 1278 32 US-10-208-731-2 Sequence 5494, Ap
62 2402.5 34.8 1278 37 US-10-218-140-5494 Sequence 1542, Ap
63 2402.5 34.8 1278 37 US-10-741-600-1542 Sequence 530, App
64 2402.5 34.8 1278 37 US-10-741-601-530 Sequence 4324, App
65 2402.5 34.8 1278 37 US-10-756-149-4324 Sequence 1261, App
66 2402.5 34.8 1278 37 US-10-796-280-1261 Sequence 810, App
67 2402.5 34.8 1278 37 US-10-796-307-810 Sequence 949, App
68 2402.5 34.8 1278 38 US-10-821-801-949 Sequence 952, App
69 2402.5 34.8 1278 39 US-10-995-561-952 Sequence 1078, Ap
70 2402.5 34.8 1278 40 US-11-066-316A-1078 Sequence 129, App
71 2402.5 34.8 1278 41 US-11-191-274-129 Sequence 130, App
72 2402.5 34.8 1278 41 US-11-191-274-130 Sequence 141, App
73 2402.5 34.8 1278 42 US-11-222-021-441 Sequence 148, App
74 2402.5 34.8 1278 48 US-60-423-552-148 Sequence 148, App
75 2402.5 34.8 1278 48 US-60-427-579-148 Sequence 1424, A
76 2402.5 34.8 1278 48 US-60-453-050-14124 Sequence 14124, A
77 2402.5 34.8 1278 48 US-60-453-135-14124 Sequence 14124, A
78 2402.5 34.8 1278 48 US-60-466-412-14124 Sequence 700, App
79 2402.5 34.8 1278 49 US-60-524-882-700 Sequence 476, App
80 2402.5 34.8 1278 49 US-60-568-219-476 Sequence 129, App
81 2402.5 34.8 1278 49 US-60-592-195-129 Sequence 130, App
82 2402.5 34.8 1278 50 US-60-592-195-130 Sequence 137, App
83 2402.5 34.8 1278 50 US-60-651-235-197 Sequence 138, App
84 2402.5 34.8 1278 50 US-60-651-235-198 Sequence 138, App
85 2402.5 34.8 1278 50 US-60-680-002-1038 Sequence 191, App
86 2402.5 34.8 1278 51 US-60-720-274-191 Sequence 440, App
87 2402.5 34.8 1279 42 US-11-223-021-440 Sequence 1037, Ap
88 2385 34.5 1319 32 US-10-208-731-4 Sequence 4, Appl1
89 2262 32.7 1287 1 PCT-US03-24982A-291 Sequence 291, App
90 2262 32.7 1287 26 US-09-614-150-12003 Sequence 12003, A
91 2262 32.7 1287 26 US-09-614-150A-12003 Sequence 12003, A
92 2262 32.7 1287 40 US-11-097-143-12003 Sequence 274, App
93 2087 30.2 1525 46 US-60-229-511-274 Sequence 129, App
94 2083 30.1 1529 46 US-60-212-657-129 Sequence 12035, A
95 2036.5 29.5 1157 45 US-60-191-637-12035 Sequence 53619, A
96 1898.5 27.5 1279 34 US-10-449-902-53619 Sequence 2679, Ap
97 1864.5 27.0 1223 26 US-09-614-150-2679 Sequence 2679, Ap
98 1864.5 27.0 1223 26 US-09-614-150A-2679 Sequence 2679, Ap
99 1864.5 27.0 1223 40 US-11-097-143-2679 Sequence 514, App
100 1864.5 27.0 1223 45 US-60-161-932-514 Sequence 514, App

ALIGNMENTS

RESULT 1
PCT-US03-22467-4
; Sequence 4, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-22467-4

Query Match 100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MAEAGLRLWALLRLAQSEPTTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60

Db QY
1 MAEAGLRLWALLRLAQSEPTTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
61 TPARKITGDHLLILLOKICPRLYTGNTACCSAKOLVSEASLSITKALLTRCPACSNF 120
61 TPARKITGDHLLILLOKICPRLYTGNTACCSAKOLVSEASLSITKALLTRCPACSNF 120
121 VNLCHNTCSNPQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180
121 VNLCHNTCSNPQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180
181 ATLAVGTWCGYVGSALCNAQRLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQLNEGV 240
181 ATLAVGTWCGYVGSALCNAQRLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQLNEGV 240
241 ARCHESQDDVATSCODCAAPALPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
241 ARCHESQDDVATSCQDCAAPALPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWVASWPLTILVLSV 360
301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWVASWPLTILVLSV 360
361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHQHFGPPFRTRNQVILTAPNRSY 420
361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHQHFGPPFRTRNQVILTAPNRSY 420
421 RYDSLILGPKNFGSLDLDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
421 RYDSLILGPKNFGSLDLDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVMDKHDFLYCANAPLTFKDGTLAL 540
481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVMDKHDFLYCANAPLTFKDGTLAL 540
541 SCMDYGAVPFPFLAIGYKGYKDYSEABALIMTFLSNYPAGDPRLAQAALWEEAFLEEM 600
541 SCMDYGAVPFPFLAIGYKGYKDYSEABALIMTFLSNYPAGDPRLAQAALWEEAFLEEM 600
601 RAFQRMAGMQVTTAERSLEDEINRTTADPLIFATSYIVIFLYISIALGYSWSRV 660
601 RAFQRMAGMQVTTAERSLEDEINRTTADPLIFATSYIVIFLYISIALGYSWSRV 660
661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSGADNIFVLE 720
661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSGADNIFVLE 720
721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTPALTSLGLAV 780
721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTPALTSLGLAV 780
781 ILDFLLQMSAFVALLSLSKQESRLDVCCCKPQELPPPCGGGLLGFQKAYAPFL 840
781 ILDFLLQMSAFVALLSLSKQESRLDVCCCKPQELPPPCGGGLLGFQKAYAPFL 840
841 LHWITRGVVLLFLFALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
841 LHWITRGVVLLFLFALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
901 YFVTTLGNFSSSEAGNNAICSSAGCNFNFSFTQKIQYATEFPFQSYLAIPASSWDDFIDW 960
901 YFVTTLGNFSSSEAGNNAICSSAGCNFNFSFTQKIQYATEFPFQSYLAIPASSWDDFIDW 960
961 LTPSSCCRLYISGPNKDKFCPTVNSLCLKNCMSITMGSVRPSVEQPHKYLPPWLNDRP 1020
961 LTPSSCCRLYISGPNKDKFCPTVNSLCLKNCMSITMGSVRPSVEQPHKYLPPWLNDRP 1020
1021 NTKCPKGGLAAYSTSVNLTSDQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
1021 NTKCPKGGLAAYSTSVNLTSDQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
1081 RKVPGTDPAFVFPFVPTTNNVFYEQVITILPEGLFMLSLCLVPTFAVSCILLGLDLSGLL 1140
1081 RKVPGTDPAFVFPFVPTTNNVFYEQVITILPEGLFMLSLCLVPTFAVSCILLGLDLSGLL 1140


```
; FILE REFERENCE: 36134-PCT 074669, 0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US 60/537,341
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01469-4

Query Match      100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSSGLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLLOKICPRLYTGNTQACCSAKQLVSLASLSTKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLLOKICPRLYTGNTQACCSAKQLVSLASLSTKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVPA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVPA 180
Qy 181 ATLAGVTGCVYGSALCNAQRLNFGDGTGNGLAFLDITFHLLEPGQVAGSGIQPLNEGV 240
Db 181 ATLAGVTGCVYGSALCNAQRLNFGDGTGNGLAFLDITFHLLEPGQVAGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATSCDQCAASCPAIRPAQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATSCDQCAASCPAIRPAQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGRVAPARKSKMVDPKGTSLSDKLSFSTHLLGQFFQCGWTWASWPLTILVLSV 360
Db 301 LLVGRVAPARKSKMVDPKGTSLSDKLSFSTHLLGQFFQCGWTWASWPLTILVLSV 360
Qy 361 IPWVLAAGLVTELTDDPVELWSAPNSQAREKAFHDHGFPPFRTNQVILTAENRSY 420
Db 361 IPWVLAAGLVTELTDDPVELWSAPNSQAREKAFHDHGFPPFRTNQVILTAENRSY 420
Qy 421 RYDSSLGPKNFGSGLDLDLLELEQLERLRLQVWSPQAQNLISLQDICVAPLNPDNT 480
Db 421 RYDSSLGPKNFGSGLDLDLLELEQLERLRLQVWSPQAQNLISLQDICVAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDHLFCANAPLTFCQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDWKDHLFCANAPLTFCQGTALAL 540
Qy 541 SCWADYGAPVFPFLAIGGKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Qy 601 RAFQRRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIIVIFLYISLALGSYSWSRV 660
Db 601 RAFQRRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIIVIFLYISLALGSYSWSRV 660
Qy 661 MVDKATLGLGVAVVILGAVMAAMGFFSYLGRIRSSLVILQVPPFLVLSVGAADNIFIFVLE 720
Db 661 MVDKATLGLGVAVVILGAVMAAMGFFSYLGRIRSSLVILQVPPFLVLSVGAADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTTPMPAVRTTALTSGLAV 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTTPMPAVRTTALTSGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQBASRLDVCCCKVQDELPPPGQEGLLIGFFQKAVAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQBASRLDVCCCKVQDELPPPGQEGLLIGFFQKAVAPFL 840

RESULT 4
US-10-621-758A-4
; Sequence 4, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NEC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4

Query Match      100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSSGLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLLOKICPRLYTGNTQACCSAKQLVSLASLSTKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLLOKICPRLYTGNTQACCSAKQLVSLASLSTKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVPA 180
```

Db 121 VNLHCHTSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAQSQSDSCSRVRPAA 180
QY 181 ATLAGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQOAVGSGIOLNEGV 240
Db 181 ATLAGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQOAVGSGIOLNEGV 240
QY 241 ARCNSQGGDDVATCSCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNSQGGDDVATCSCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVRVAPARDKSKWDPKGTSLSDKLSFSTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKWDPKGTSLSDKLSFSTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTPAPNRSY 420
Db 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTPAPNRSY 420
QY 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQWSPQAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQWSPQAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHPLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHPLYCANAPLTPKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLMEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLMEAFLEEM 600
QY 601 RAPQRMAGHFWQVTFATRSLEDEINRTTAEDLPFATSYIVIFLYISALGYSWSRV 660
Db 601 RAPQRMAGHFWQVTFATRSLEDEINRTTAEDLPFATSYIVIFLYISALGYSWSRV 660
QY 661 MVDKATLGGVAVVIGVAVMAAGFYSYLGIRSSIVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MVDKATLGGVAVVIGVAVMAAGFYSYLGIRSSIVILQVVPFLVLSVGDNIPIFVLE 720
QY 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEALCFFLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEALCFFLGALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQBSRLDVCCKVKPOELPPPGQEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQBSRLDVCCKVKPOELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHMWITRGVULLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
Db 841 LHMWITRGVULLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
QY 901 YFVTTILGYNFSBAGMNAICSSAGCNPFSTQKIQVATFPQSYLAIIPASSWVDDFDIW 960
Db 901 YFVTTILGYNFSBAGMNAICSSAGCNPFSTQKIQVATFPQSYLAIIPASSWVDDFDIW 960
QY 961 LTPSSCCRLYIYSGPNKDXFCPSVTNSLCLKNCMSITMGSRVPSVQPHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYIYSGPNKDXFCPSVTNSLCLKNCMSITMGSRVPSVQPHKYLFWFLNDRP 1020
QY 1021 NIKCPKGGAAVSTSNLSDGOVLASRPMWYHKPLKNSQDYTEALRAARELAANTADL 1080
Db 1021 NIKCPKGGAAVSTSNLSDGOVLASRPMWYHKPLKNSQDYTEALRAARELAANTADL 1080
QY 1081 RKVPGTDPAPFPPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGGLL 1140
Db 1081 RKVPGTDPAPFPPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGGLL 1140
QY 1141 NLLSIVMILVDTVGFMAWDIISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGFMAWDIISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVPAVAMTNLPGLVGLAKAQLOIIPFRNLNLTITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVPAVAMTNLPGLVGLAKAQLOIIPFRNLNLTITLGLLHGLVFLPV 1260

QY 1261 ILSYVGPDVNPALALEQKRAEBAVAAMVASCPNHPFSRVSTADNIYVNHSPFGSIKAGA 1320
Db 1261 ILSYVGPDVNPALALEQKRAEBAVAAMVASCPNHPFSRVSTADNIYVNHSPFGSIKAGA 1320
QY 1321 ISNFLPNNGROP 1332
Db 1321 ISNFLPNNGROP 1332

RESULT 5

US-10-646-301A-4
; Sequence 4, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-4

Query Match 100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Fred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTHOPGYCAFYDECGKPELGSGLMTLSNVCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTHOPGYCAFYDECGKPELGSGLMTLSNVCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHTSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAQSQSDSCSRVRPAA 180
Db 121 VNLHCHTSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAQSQSDSCSRVRPAA 180
QY 181 ATLAGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQOAVGSGIOLNEGV 240
Db 181 ATLAGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQOAVGSGIOLNEGV 240
QY 241 ARCNSQGGDDVATCSCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNSQGGDDVATCSCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVRVAPARDKSKWDPKGTSLSDKLSFSTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKWDPKGTSLSDKLSFSTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTPAPNRSY 420
Db 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTPAPNRSY 420
QY 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQWSPQAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQWSPQAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHPLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHPLYCANAPLTPKDGFTALAL 540

QY 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALINTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
DB 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALINTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
QY 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
DB 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
QY 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSVGDADNIFIFVLE 720
DB 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSVGDADNIFIFVLE 720
QY 721 YQRLPRRPGEPREHVGRAIGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREHVGRAIGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
QY 841 LHMWITRGVVLLFLALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHMWITRGVVLLFLALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCANNFSFTOKI QYATEFPQSVLAIPASSWVDDFIDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCANNFSFTOKI QYATEFPQSVLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFINDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFINDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVPPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
DB 1081 RKVPGTDPAPFVPPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
QY 1141 NLSIWMILVDTVGFMAWDISYNAVSLNLYSAVGMVSEFVSHIRSPAISTKPTWLER 1200
DB 1141 NLSIWMILVDTVGFMAWDISYNAVSLNLYSAVGMVSEFVSHIRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRMLLTITLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRMLLTITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVNPALALEQKRAEEAAVAVMVASCPNHSRVSSTADNIYVNHSPFGSIKGAGA 1320
DB 1261 ILSYVGPVNPALALEQKRAEEAAVAVMVASCPNHSRVSSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGROF 1332
DB 1321 ISNFLPNNGROF 1332

RESULT 6

US-10-663-208A-4
; Sequence 4, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J061603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; PRIOR FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758

; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-663-208A-4

Query Match 100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLLAQSEPYTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINVTRVAOLGAGQLPAVVAZEAFYOHSPABQSYDSCSRVRVPA 180
DB 121 VNLHCHNTCSNQSLFINVTRVAOLGAGQLPAVVAZEAFYOHSPABQSYDSCSRVRVPA 180
QY 181 ATLAGVTCGVYGSALCNAQRLNFOQDGTGNGLAPLDITTEHLEPGQAVGSGIQPLNEGV 240
DB 181 ATLAGVTCGVYGSALCNAQRLNFOQDGTGNGLAPLDITTEHLEPGQAVGSGIQPLNEGV 240
QY 241 ARCHESQDDVATCSCQCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCHESQDDVATCSCQCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSNQVDPKKGTSLSDKLSFSTHTLLQGFQGWGTWASWPLTILVLSV 360
DB 301 LLVGFRVAPARDKSNQVDPKKGTSLSDKLSFSTHTLLQGFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHGFPPFTNQVILTAENRSY 420
DB 361 IPVVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHGFPPFTNQVILTAENRSY 420
QY 421 RYDSELLGPKNPSGILDLDLLELELQERLHLQVMSPEAQRNISLQDICYAPLNPDNT 480
DB 421 RYDSELLGPKNPSGILDLDLLELELQERLHLQVMSPEAQRNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540
QY 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALINTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
DB 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALINTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
QY 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
DB 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
QY 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSVGDADNIFIFVLE 720
DB 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSVGDADNIFIFVLE 720
QY 721 YQRLPRRPGEPREHVGRAIGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREHVGRAIGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGFQKAYAPFL 840
QY 841 LHMWITRGVVLLFLALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPV 900

Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFANDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFANDRP 1020
QY 1021 NIKCPKGGAAVSTSNLTSQGVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGGAAVSTSNLTSQGVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPCTDPAFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPCTDPAFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLTITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLTITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVDPNPALEQKAEAEAAVAVMVASCPNHPSTRYSTADNIYNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPVDPNPALEQKAEAEAAVAVMVASCPNHPSTRYSTADNIYNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGROF 1332
Db 1321 ISNFLPNNGROF 1332

RESULT 7

US-10-736-769-4
; Sequence 4, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-4

Query Match 100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLAQSPEYTTIHOPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSPEYTTIHOPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLQKICPRLYTGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120

Db 61 TPARKITGDHLILLQKICPRLYTGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFARQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFARQSYDSCSRVRVPA 180
QY 181 ATLAIVGTWCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGSIQPLNEGV 240
Db 181 ATLAIVGTWCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGSIQPLNEGV 240
QY 241 ARCNSQGGDDVATCQDCCAASCPAARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGGDDVATCQDCCAASCPAARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
Db 301 LLVGFRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVPTTELTPDVELWSAPNSQARSEKAFHDQHFGPFPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVPTTELTPDVELWSAPNSQARSEKAFHDQHFGPFPFRTNQVILTAPNRSSY 420
QY 421 RYDSILLGPKNPSGILDLDLLELELERLHLQVMSPEAQRNLSLODICYAPLNPDT 480
Db 421 RYDSILLGPKNPSGILDLDLLELELERLHLQVMSPEAQRNLSLODICYAPLNPDT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTQSDVKDHFYCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTQSDVKDHFYCANAPLTFKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAPQRMAGMQVTTAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALGYSYWSRV 660
Db 601 RAPQRMAGMQVTTAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALGYSYWSRV 660
QY 661 MYDSKATILGLGVAVVLGAVMAAMGFFSYGIRSSILVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAMGFFSYGIRSSILVILQVVPFLVLSVGDNIPIFVLE 720
QY 721 YORLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACIPFLGALTTPMPAVRTTALTSLGLAV 780
Db 721 YORLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACIPFLGALTTPMPAVRTTALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLSKQESRLDVCCCKQBELPPPCQGGELLGPFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLSKQESRLDVCCCKQBELPPPCQGGELLGPFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFANDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFANDRP 1020
QY 1021 NIKCPKGGAAVSTSNLTSQGVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGGAAVSTSNLTSQGVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPCTDPAFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPCTDPAFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200

```
QY 1201 AKEATISMSAVFAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLLITLLGLLHGLVFLPV 1260
DB 1201 AKEATISMSAVFAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLLITLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPRSRVSTADNIYNHSPGSIKGAGA 1320
DB 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPRSRVSTADNIYNHSPGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 8
US-10-750-386-4
; Sequence 4, Application US/10750386
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; APPLICANT: Chapman, Kevin
; APPLICANT: Goulet, Mark
; APPLICANT: Ujjainwalla, Feroze
; APPLICANT: Altmann, Scott W
; APPLICANT: Davis, Chip
; APPLICANT: Bull, Herb
; APPLICANT: Thornberry, Nancy A
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONISTS
; FILE REFERENCE: A36104 074669.0103
; CURRENT APPLICATION NUMBER: US/10/750,386
; CURRENT FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-750-386-4

Query Match 100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLGRWLLWALLRLAQSEPYTTHQPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLGRWLLWALLRLAQSEPYTTHQPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLASLITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLASLITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNOSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDSCSRVRVPA 180
DB 121 VNLHCHNTCSPNOSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDSCSRVRVPA 180
QY 181 ATLAVGTMCGYGSALCNQARWLNFGDGTGNGLAPLDTIFHLLEPGQAVGSGIQLNEGV 240
DB 181 ATLAVGTMCGYGSALCNQARWLNFGDGTGNGLAPLDTIFHLLEPGQAVGSGIQLNEGV 240
QY 241 ARCNEGQDDVATCSQDCAASCPAARPOALDSTFYLQMPGSLVLIILCSVFVAVVTI 300
DB 241 ARCNEGQDDVATCSQDCAASCPAARPOALDSTFYLQMPGSLVLIILCSVFVAVVTI 300
QY 301 LLVGFVRVAPARDKSMVDPKKTSLSKLSFSFTHLLGQPFQGWGTWASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSMVDPKKTSLSKLSFSFTHLLGQPFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLETTDPVELSAPNSQARSSEKAFHDQHFPGFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTLETTDPVELSAPNSQARSSEKAFHDQHFPGFPFRTNQVILTAPNRSY 420
QY 421 RYDSLLIGPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPNDT 480
DB 421 RYDSLLIGPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPNDT 480
```

```
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGTTALAL 540
QY 541 SCMADYCAPVPPFLAIGGYKGKDYSEABALIMTFSLNNYPAGDPRLAQAOKLWBEAFLEEM 600
DB 541 SCMADYCAPVPPFLAIGGYKGKDYSEABALIMTFSLNNYPAGDPRLAQAOKLWBEAFLEEM 600
QY 601 RAFQRRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYTIVIFLYTSLAGSYSSMSRV 660
DB 601 RAFQRRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYTIVIFLYTSLAGSYSSMSRV 660
QY 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVSGADNIFIFVLE 720
DB 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVSGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHIGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREVIHIGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSLGLAV 780
QY 781 ILDFELLOMSAFVALLSLDSKQEBASRLDVCCKVQPOELPPPGQEGGLLGFQKAYAPFL 840
DB 781 ILDFELLOMSAFVALLSLDSKQEBASRLDVCCKVQPOELPPPGQEGGLLGFQKAYAPFL 840
QY 841 LHMWTRGVVLLFLALFGVSLYSMCHISVGLDQBLALPKOSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHMWTRGVVLLFLALFGVSLYSMCHISVGLDQBLALPKOSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNPNFSFTQKIQYATERPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNPNFSFTQKIQYATERPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQEFHKLVPFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQEFHKLVPFLNDRP 1020
QY 1021 NIKCPKGGLAAYSTVNLTSDGQVLASRFMAHYHKPKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGGLAAYSTVNLTSDGQVLASRFMAHYHKPKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVFPYTIITNVFYEQLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL 1140
DB 1081 RKVPGTDPAPFVFPYTIITNVFYEQLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL 1140
QY 1141 NLLSVMLVILVDVTFGPMALWDISYNAVSLINLVASVGMSEVFPVSHITRGSFAISTKPTWLER 1200
DB 1141 NLLSVMLVILVDVTFGPMALWDISYNAVSLINLVASVGMSEVFPVSHITRGSFAISTKPTWLER 1200
QY 1201 AKEATISMSAVFAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLLITLLGLLHGLVFLPV 1260
DB 1201 AKEATISMSAVFAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLLITLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPRSRVSTADNIYNHSPGSIKGAGA 1320
DB 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPRSRVSTADNIYNHSPGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 9
PCT-US01-04098A-1831
; Sequence 1831, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
```

;; PRIOR APPLICATION NUMBER: 09/728,422
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: 09/693,325
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 09/663,561
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: 09/654,936
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 09/620,325
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: 09/598,075
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 3960
;; SOFTWARE: Custom
;; SEQ ID NO 1831
;; LENGTH: 1332
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US01-04098A-1831

Query Match 99.8%; Score 6896; DB 1; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MABAGLRGWLWALLRLAQSPYTTIHOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MABAGLRGWLWALLRLAQSPYTTIHOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVLEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSNQSIFINTRVAQAGAGQLPVAVYAEFYQHSFAEQSDSCSRVRVPA 180
DB 121 VNLHCHTCSNQSIFINTRVAQAGAGQLPVAVYAEFYQHSFAEQSDSCSRVRVPA 180
QY 181 ATLAVGTGCVGVSALCNAQWLNFGDGTGNGLAPLDTFHLLEPQAVGSGIQLNEGV 240
DB 181 ATLAVGTGCVGVSALCNAQWLNFGDGTGNGLAPLDTFHLLEPQAVGSGIQLNEGV 240
QY 241 ARCNESSQGDVATCSQDCAACSPALARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESSQGDVATCSQDCAACSPALARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFPGWGTWVASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFPGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDOHFGFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDOHFGFPFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELLELQERLHLQWVSPAQRNLSLODICVAPLNPDNT 480
DB 421 RYDLSLLGPKNFGSLDLDLLELLELQERLHLQWVSPAQRNLSLODICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTMGQTSQVDWKDHFLYCANAPLTFKDGATLAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTMGQTSQVDWKDHFLYCANAPLTFKDGATLAL 540
QY 541 SCWADYGAPVFPFLAIGGKGYGEAEALIMTFSLNYPAGDPPLAQKWEAEFLPEM 600
DB 541 SCWADYGAPVFPFLAIGGKGYGEAEALIMTFSLNYPAGDPPLAQKWEAEFLPEM 600
QY 601 RAFQRMAGMFOVTTFAERSLDEINRTTAEDLPFATSYIVIFLYISLALGYSWSGRV 660
DB 601 RAFQRMAGMFOVTTFAERSLDEINRTTAEDLPFATSYIVIFLYISLALGYSWSGRV 660
QY 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLIGIRSSLVILQVVPFLVLSVGADNIFVLE 720

DB 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLIGIRSSLVILQVVPFLVLSVGADNIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACPLGALTMPAVRTTALTSGLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACPLGALTMPAVRTTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKVQBELPPPGQEGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCCKVQBELPPPGQEGGLLGFQKAYAPFL 840
QY 841 LHMITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHMITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVVTLLGNFSSSEAGNNAICSSAGCNPNFSTQKIYATFPPQSYLAIPASSWVDDPIDW 960
DB 901 YFVVTLLGNFSSSEAGNNAICSSAGCNPNFSTQKIYATFPPQSYLAIPASSWVDDPIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQEFHKYLPWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQEFHKYLPWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQVTLASRPMAYHKLPKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTVNLTSQVTLASRPMAYHKLPKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFVPPYTTITNVFYEQYLTILPEGLFMLCLVPTPAVSCLLGLDLSGGL 1140
DB 1081 RKVPGTDPAFVPPYTTITNVFYEQYLTILPEGLFMLCLVPTPAVSCLLGLDLSGGL 1140
QY 1141 NLLSVIMILVDVTGFMALWDISYNAVSLINLVSAYGMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSVIMILVDVTGFMALWDISYNAVSLINLVSAYGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGIILVGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
DB 1201 AKEATISGMSAVFAGVAMTNLPGIILVGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIVVNHSPGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIVVNHSPGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 10
US-10-239-316-9
; Sequence 9, Application US/10239316
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-9

Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAEAGRLWLLWALLRLAQSEPYTTIHQGYCAFYDECCKNPGLSGMLTSLNVCSLSN 60
Db 1 MAEAGRLWLLWALLRLAQSEPYTTIHQGYCAFYDECCKNPGLSGMLTSLNVCSLSN 60
Qy 61 TPARKITGDHILLILLOKICPRLYTGNTQACCSAQQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILLILLOKICPRLYTGNTQACCSAQQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPAA 180
Db 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPAA 180
Qy 181 ATLVGTCGVYGSALCNAQRWLPFGDTGNGLAPLDTFHLLBPQAVGSGIQPLNEGV 240
Db 181 ATLVGTCGVYGSALCNAQRWLPFGDTGNGLAPLDTFHLLBPQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSQDCQCAAPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCQCAAPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLVS 360
Db 301 LLVGRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLVS 360
Qy 361 IPVVALAAGLVFTELTTPDELWLSAPNSQARSEKAFHQHCPFRFTNQVILTAENRSSY 420
Db 361 IPVVALAAGLVFTELTTPDELWLSAPNSQARSEKAFHQHCPFRFTNQVILTAENRSSY 420
Qy 421 RYDSLGLGPNFSGILDLDLLELLELOERLHLQVMSPEAQRNISLODICVAPLNPDNT 480
Db 421 RYDSLGLGPNFSGILDLDLLELLELOERLHLQVMSPEAQRNISLODICVAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKHDFLYCANAPLTPKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKHDFLYCANAPLTPKQGTALAL 540
Qy 541 SCHADYGAVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLREM 600
Db 541 SCHADYGAVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLREM 600
Qy 601 RAFORMMAGMFWTTFABRSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSWSRV 660
Db 601 RAFORMMAGMFWTTFABRSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSWSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAMGPFYSYLGIRSSLVILQVVPFLVSVGADNIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAMGPFYSYLGIRSSLVILQVVPFLVSVGADNIFVLE 720
Qy 721 YQRLPRPGEPREVHIGRALGRVAPSMILCSISEATCFPLGALTMPAVRTALTSLGLAV 780
Db 721 YQRLPRPGEPREVHIGRALGRVAPSMILCSISEATCFPLGALTMPAVRTALTSLGLAV 780
Qy 781 ILDFLLQMAFAVALLSLDKRQASRLDCCCVPQELPPPGQEGGLLGFQKAYAPFL 840
Db 781 ILDFLLQMAFAVALLSLDKRQASRLDCCCVPQELPPPGQEGGLLGFQKAYAPFL 840
Qy 841 LHMITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHMITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTOKIQYATEPPEOSYLAIPASSWVDDPIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTOKIQYATEPPEOSYLAIPASSWVDDPIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRSPVBFHKYLPWFNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRSPVBFHKYLPWFNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAVYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAVYHKPLKNSQDYTEALRAARELANITADL 1080

RESULT 11

US-10-258-899A-1831
; Sequence 1831, Application US/10258899A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Cao, Yicheng
; APPLICANT: Ma, Yunquing
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Xue, Aidong
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Goodrich, Ryle
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 787CIP2-2G/US
; CURRENT APPLICATION NUMBER: US/10/258,899A
; PRIOR FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: PCT/US01/04098
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/774,434
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/554,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT

Qy 1081 RKVGTDPAPAFEVFPYTTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLILGLDLRSGLL 1140
Db 1081 RKVGTDPAPAFEVFPYTTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLILGLDLRSGLL 1140
Qy 1141 NLLSVIMILVDVTGFMALWDISYNAVSLINIVSAGMSVBFVSHITSPASTKPTWLER 1200
Db 1141 NLLSVIMILVDVTGFMALWGISYNAVSLINIVSAGMSVBFVSHITSPASTKPTWLER 1200
Qy 1201 AKRATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFFFRLNLLITLLGLLHGLVFLPV 1260
Db 1201 AKRATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFFFRLNLLITLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGDDVNPALALEQKRAEEAAVAAVMVASCNPNHPSRVSTADNIIYNHSPGSIKGAGA 1320
Db 1261 ILSYVGDDVNPALALEQKRAEEAAVAAVMVASCNPNHPSRVSTADNIIYNHSPGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

```

; ORGANISM: Homo sapiens
US-10-258-899A-1831

Query Match      99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKI CRLTYTGNTQACCSAKQLVSEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKI CRLTYTGNTQACCSAKQLVSEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTSPNQSFLINTRVAQLGAGQLPAVAYEAFYQHSFARQSYDSCSRVRVPA 180
DB 121 VNLHCHNTSPNQSFLINTRVAQLGAGQLPAVAYEAFYQHSFARQSYDSCSRVRVPA 180
QY 181 ATLAGVTMGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
DB 181 ATLAGVTMGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
QY 241 ARCNSQGDVATCSCODCAASCPTARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNSQGDVATCSCODCAASCPTARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
DB 301 LLVGFRVAPARDKSMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDTPVELWSAPNSQARSEKAFHQHGFPPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTDTPVELWSAPNSQARSEKAFHQHGFPPFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPNKSGILDLDLLELLELQRLHLQWSPEAQRNLSLQICVAPLNPDNT 480
DB 421 RYDLSLLGPNKSGILDLDLLELLELQRLHLQWSPEAQRNLSLQICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDFLYCANAPLTPKDGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDFLYCANAPLTPKDGTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKWEBAFLEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKWEBAFLEM 600
QY 601 RAPORMAGMFOVTTAERSLDEINRTAEDLPATSYIVIFLYISALGSYSWSRV 660
DB 601 RAPORMAGMFOVTTAERSLDEINRTAEDLPATSYIVIFLYISALGSYSWSRV 660
QY 661 MYDSKATLGLGVAVVLGNVMAAMGFSSYLGRSSILVILQVVPFLVLSVGDNIIFVLE 720
DB 661 MYDSKATLGLGVAVVLGNVMAAMGFSSYLGRSSILVILQVVPFLVLSVGDNIIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTTPMPAVRTEALTSLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTTPMPAVRTEALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCKVKQELPPPPQCGEGLLIGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCKVKQELPPPPQCGEGLLIGFFQKAYAPFL 840
QY 841 LHWITRGVWVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPV 900
DB 841 LHWITRGVWVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPV 900
QY 901 YPVTTLGNVFSSEAGNNAICSSAGCNFSFTOKIQTATEFFPQSXYLAIPASSWVDQFDW 960
DB 901 YPVTTLGNVFSSEAGNNAICSSAGCNFSFTOKIQTATEFFPQSXYLAIPASSWVDQFDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLFWFLNDRP 1020

; Sequence 1831, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1831

Query Match      99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKI CRLTYTGNTQACCSAKQLVSEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKI CRLTYTGNTQACCSAKQLVSEASISITKALLTRCPACSDNF 120
```

Db 61 TPARKITGDHLLILLQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAGVTCGCVGYSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEG 240
Db 181 ATLAGVTCGCVGYSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEG 240
Qy 241 ARCNEQGDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFVVTI 300
Qy 301 LLVGRVAPADKSKWDPKKGTSLSKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPADKSKWDPKKGTSLSKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHDOHFGPPRTNQVILTAENRSY 420
Db 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHDOHFGPPRTNQVILTAENRSY 420
Qy 421 RYDLSLLGPKNPSGLDLDLLELEQLERLHLQWSPQARQNSLQDICYAPLNPNT 480
Db 421 RYDLSLLGPKNPSGLDLDLLELEQLERLHLQWSPQARQNSLQDICYAPLNPNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTFKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTFKQGTALAL 540
Qy 541 SCMADYGAPVFPFLAIGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAKWEAEFLSEM 600
Db 541 SCMADYGAPVFPFLAIGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAKWEAEFLSEM 600
Qy 601 RAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSYSWRSV 660
Db 601 RAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSYSWRSV 660
Qy 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLSVLQVVPFLVSVGADNIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLSVLQVVPFLVSVGADNIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSSEALCFPLGALTTPMPAVRTEALTSGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSSEALCFPLGALTTPMPAVRTEALTSGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKVQELPPPPQCGEGLLIGFPQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKVQELPPPPQCGEGLLIGFPQKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNATCSSAGCNNSFTOKI QYATEFPQSUYLAIPASSWDDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNATCSSAGCNNSFTOKI QYATEFPQSUYLAIPASSWDDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCILKNCMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCILKNCMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAFEPPFYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDURSGLL 1140
Db 1081 RKVPGTDPAFEPPFYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDURSGLL 1140
Qy 1141 NLLSIWMILVDTVGFNALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIWMILVDTVGFNALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200

RESULT 13

US-11-218-141-1831
; Sequence 1831, Application US/11218141
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/11/218,141
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-218-141-1831

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLLAQSEPYTIIHQPGYCAFVDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTIIHQPGYCAFVDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLILLQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAGVTCGCVGYSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEG 240
Db 181 ATLAGVTCGCVGYSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEG 240
Qy 241 ARCNEQGDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFVVTI 300

QY 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSV 420
DB 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSV 420
QY 421 RYDSSLGPKNFGSILDLLOLLLELLELQERLRLHQLWSPEAQRNLSLQDICYAPLNPNT 480
DB 421 RYDSSLGPKNFGSILDLLOLLLELLELQERLRLHQLWSPEAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGXTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGXTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
QY 601 RAFQRMAGMFQVTFWAEERLDEINRTTAEDLPIFATSYIVIFLYISALAGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTFWAEERLDEINRTTAEDLPIFATSYIVIFLYISALAGSYSSWSRV 660
QY 661 MVDSKATLGGVAVVVLGAVMAAGPFSYLGIRSSLVILQVVPFLVSVGADNIFIFVLE 720
DB 661 MVDSKATLGGVAVVVLGAVMAAGPFSYLGIRSSLVILQVVPFLVSVGADNIFIFVLE 720
QY 721 YQRLPRRPEPREVHIGRALGRVAPSMJLCSSEAI CFFLGALTPMPAVRTFALTSGLA 780
DB 721 YQRLPRRPEPREVHIGRALGRVAPSMJLCSSEAI CFFLGALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQBSRLDCCVCKPOELPPPGQEGLLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQBSRLDCCVCKPOELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHI SVGLDQELALPKDSYLLDYFLFNRYFEVGA 900
DB 841 LHWITRGVLLFLALFGVLSYMSCHI SVGLDQELALPKDSYLLDYFLFNRYFEVGA 900
QY 901 YFVTTLYGNSFSEAGNNAICSSAGCNPFSTQKI QYATFEPQSYLAI PASSWDDFD 960
DB 901 YFVTTLYGNSFSEAGNNAICSSAGCNPFSTQKI QYATFEPQSYLAI PASSWDDFD 960
QY 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSITMGSVRPSVEQPHKYLFPWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSITMGSVRPSVEQPHKYLFPWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGQVLAASFMAVYHKLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDGQVLAASFMAVYHKLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEPVPPYITINVTQYQYTLILPEGLFMLSCLVPTFAVSCLLGLDRLSGLL 1140
DB 1081 RKVPGTDPAFEPVPPYITINVTQYQYTLILPEGLFMLSCLVPTFAVSCLLGLDRLSGLL 1140
QY 1141 NLLSIWMILVDVTGFMALWDISVNASLNLNLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSIWMILVDVTGFMALWDISVNASLNLNLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMGSAVAGVAMTNLPGLVGLAKAQLIQIFPFRNLNLITLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVAGVAMTNLPGLVGLAKAQLIQIFPFRNLNLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPNPALALEQKAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSINGAGA 1320
DB 1261 ILSYVGPDPNPALALEQKAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSINGAGA 1320
QY 1321 ISNLPNNNGROF 1332
DB 1321 ISNLPNNNGROF 1332

RESULT 14

US-11-242-459-9
; Sequence 9, Application US/11242459
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-9

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGLMTLSNVSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGLMTLSNVSLSN 60
QY 61 TPARKITGDLHLLLOKICPRLYTGPNTOACCSAKQLVLSIASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDLHLLLOKICPRLYTGPNTOACCSAKQLVLSIASISITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVYAEAFYQHSFAEQSDSCSRVPA 180
DB 121 VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVYAEAFYQHSFAEQSDSCSRVPA 180
QY 181 ATLVGTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVSGSIQPLNEGV 240
DB 181 ATLVGTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVSGSIQPLNEGV 240
QY 241 ARCNESQGDVATCSCODCAASCPAARQALDSIFYLQMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNESQGDVATCSCODCAASCPAARQALDSIFYLQMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSV 420
DB 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSV 420
QY 421 RYDSSLGPKNFGSILDLLOLLLELLELQERLRLHQLWSPEAQRNLSLQDICYAPLNPNT 480
DB 421 RYDSSLGPKNFGSILDLLOLLLELLELQERLRLHQLWSPEAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGXTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGXTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
QY 601 RAFQRMAGMFQVTFWAEERLDEINRTTAEDLPIFATSYIVIFLYISALAGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTFWAEERLDEINRTTAEDLPIFATSYIVIFLYISALAGSYSSWSRV 660

661 MYDSKATLGLGGAVALVGNVMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLE 720
QY
Db
661 MYDSKATLGLGGAVALVGNVMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLE 720
QY
Db
721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSGLAV 780
QY
Db
721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSGLAV 780
QY
Db
781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPPGQGGGLLGFQKAYAPFL 840
QY
Db
781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPPGQGGGLLGFQKAYAPFL 840
QY
841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY
Db
841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY
901 YFVTTLGYNFSSEAGNNAICSSAGCNSFTQKIYATEFPEQSYLAIPASSWVDDFDW 960
QY
Db
901 YFVTTLGYNFSSEAGNNAICSSAGCNSFTQKIYATEFPEQSYLAIPASSWVDDFDW 960
QY
961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNKMSITWGSVRPSVEQFHKYLPWFLNDRP 1020
QY
Db
961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNKMSITWGSVRPSVEQFHKYLPWFLNDRP 1020
QY
1021 NIKCPKGGAAYSTVNLTSDQVLSRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
QY
Db
1021 NIKCPKGGAAYSTVNLTSDQVLSRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
QY
1081 RKVPGTDPAFVFPVYITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY
Db
1081 RKVPGTDPAFVFPVYITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY
1141 NLLSIVMLIVDTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY
Db
1141 NLLSIVMLIVDTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY
1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRNLNLIITLGLHGLVFLPV 1260
QY
Db
1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRNLNLIITLGLHGLVFLPV 1260
QY
1261 ILSYVGPDVNPALALEQKRAEAAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY
Db
1261 ILSYVGPDVNPALALEQKRAEAAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY
1321 ISNFLPNNGRQF 1332
QY
Db
1321 ISNFLPNNGRQF 1332

RESULT 15

PCT-US01-04098A-1830
; Sequence 1830, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; FILE REFERENCE: 21272-029
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875

; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1830

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRCWLLALLRLAQSEPYTTHQPGYCAFYDECCKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRCWLLALLRLAQSEPYTTHQPGYCAFYDECCKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNQSLFINVTRVAQIGAGQLPAVVAEYAFYQHSFABQSYDSCSRVRVPA 180
Db 121 VNLCHNTCSNQSLFINVTRVAQIGAGQLPAVVAEYAFYQHSFABQSYDSCSRVRVPA 180
QY 181 ATLAGVTCGYYGSAALCNAQRLNFOGDTGNGLAPLDTTFHLLRPGQAVGSGIQPLNEGV 240
Db 181 ATLAGVTCGYYGSAALCNAQRLNFOGDTGNGLAPLDTTFHLLRPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQDQDDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQDQDDVATCSQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSNWDPKGTSLSDKLSFSTHTLLGQFGQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSNWDPKGTSLSDKLSFSTHTLLGQFGQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFHQHFGFPFRFTQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFHQHFGFPFRFTQVILTAPNRSY 420
QY 421 RYDLSLLGPKNFGSILDLLELLELQERLHLQVMSPEAQORNTSLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSILDLLELLELQERLHLQVMSPEAQORNTSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTFCQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTFCQGTALAL 540
QY 541 SCWADYGAPVPPFLAIGGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVPPFLAIGGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFOVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSWSRV 660
Db 601 RAFQRMAGMFOVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSWSRV 660
QY 661 MYDSKATLGLGVAVVVLGNVMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVVLGNVMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSGLAV 780
Db 721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAI CFFL GALT PMPAVRTFALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPPGQGGGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPPGQGGGLLGFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900


```
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTTILGYNPSSBAGNNAICSSAGCNFSFTQKIQYATEPPEQSYLAIPASSWVDDFDW 960
Db 901 YFVTTILGYNPSSBAGNNAICSSAGCNFSFTQKIQYATEPPEQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLTVAILSPRLYSGTISAHCNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLITLPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLITLPEGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIIVMILVDTVGFMAWDIISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIIVMILVDTVGFMAWDIISYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIFFFRNLNLTLLGLLHGLVFLPVILSYVYVGPVDPNPALEQKABEAAVAVMSACP 1293
Db 1261 LIQIFFFRNLNLTLLGLLHGLVFLPVILSYVYVGPVDPNPALEQKABEAAVAVMSACP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFSGIKGAGAIISNPLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFSGIKGAGAIISNPLPNNGRQF 1359

RESULT 16
PCT-US03-22467-44
; Sequence 44, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-22467-44

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLWALLRLAQSBPYTHIHPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSBPYTHIHPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNQSFLFINTRVAOLGAGOLPAVAYEAFYQHSFABQSDSCSRVRVPA 180
Db 121 VNLHCHTCSNQSFLFINTRVAOLGAGOLPAVAYEAFYQHSFABQSDSCSRVRVPA 180
Qy 181 ATLAGVTGTCVYGSALCAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVGSGIOPINEGV 240
Db 181 ATLAGVTGTCVYGSALCAQRWLNFGDGTGNGLAPLDITFHLLEPCQAVGSGIOPINEGV 240
Qy 241 ARCNEQGGDDVATSCQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
```

```
Db 241 ARCNEQGGDDVATSCQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGPVAPARDKSKMVDPKGTSLSKLSFSFTHLLGQFFQCGMGTWASWPLTILVLSV 360
Db 301 LLVGPVAPARDKSKMVDPKGTSLSKLSFSFTHLLGQFFQCGMGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFQFFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFQFFRTNQVILTAPNRSSY 420
Qy 421 RYDLSLLGPNKPSGILDLILLLELLELRLHQLWSPEAQRNISILODIYAPLNDPNT 480
Db 421 RYDLSLLGPNKPSGILDLILLLELLELRLHQLWSPEAQRNISILODIYAPLNDPNT 480
Qy 481 SLYDCCINSLLOVFOQNRNLTLLLTANQTLMGQTSQVDMKHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLOVFOQNRNLTLLLTANQTLMGQTSQVDMKHFLYCANAPLTPKDGFTALAL 540
Qy 541 SCHADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLMEAEFLSEM 600
Db 541 SCHADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLMEAEFLSEM 600
Qy 601 RAPORMAGMPOVTTFAERSLDEINRTTAEDLPIPATSYIIVFLYISLALGSSYSWSRV 660
Db 601 RAPORMAGMPOVTTFAERSLDEINRTTAEDLPIPATSYIIVFLYISLALGSSYSWSRV 660
Qy 661 WYDSKATILGIGVAVVILGAVMAAGPPSYGLRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 WYDSKATILGIGVAVVILGAVMAAGPPSYGLRSSLVILQVVPFLVLSVGDNIPIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKQOELPPPQGEGLLIGFFOKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKQOELPPPQGEGLLIGFFOKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTTILGYNPSSBAGNNAICSSAGCNFSFTQKIQYATEPPEQSYLAIPASSWVDDFDW 960
Db 901 YFVTTILGYNPSSBAGNNAICSSAGCNFSFTQKIQYATEPPEQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLTVAILSPRLYSGTISAHCNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLITLPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLITLPEGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIIVMILVDTVGFMAWDIISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIIVMILVDTVGFMAWDIISYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIFFFRNLNLTLLGLLHGLVFLPVILSYVYVGPVDPNPALEQKABEAAVAVMSACP 1293
Db 1261 LIQIFFFRNLNLTLLGLLHGLVFLPVILSYVYVGPVDPNPALEQKABEAAVAVMSACP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFSGIKGAGAIISNPLPNNGRQF 1332
```

Db	1321	NHPSRVSTADNIYVNHSPGSIKGAGAI SNFLPNNGRQF	1359
RESULT 17			
PCT-US03-40113-44			
; Sequence 44, Application PC/TUS0340113			
; GENERAL INFORMATION:			
; APPLICANT: Schering Corporation			
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JBO1603-K3-WI			
; CURRENT APPLICATION NUMBER: PCT/US03/40113			
; CURRENT FILING DATE: 2003-12-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; PRIOR APPLICATION NUMBER: 10/663,208			
; PRIOR FILING DATE: 2003-09-16			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 44			
; LENGTH: 1359			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
PCT-US03-40113-44			
Query Match 99.5%; Score 6872.5; DB 1; Length 1359;			
Best Local Similarity 97.9%; Pred. No. 0;			
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;			
QY	1	MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKQPELSGLMTLSNVCSLSN	60
Db	1	MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKQPELSGLMTLSNVCSLSN	60
QY	61	TPARKITGDHLILLOKICRLTYGTNTQACCSAKOLVLEASLSITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLILLOKICRLTYGTNTQACCSAKOLVLEASLSITKALLTRCPACSDNF	120
QY	121	VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA	180
Db	121	VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA	180
QY	181	ATLAVGTMCVGTGSAICNAQRWLNFGDGTGNGLAFLDITFHLLEPGQAVGSIQPLNEGV	240
Db	181	ATLAVGTMCVGTGSAICNAQRWLNFGDGTGNGLAFLDITFHLLEPGQAVGSIQPLNEGV	240
QY	241	ARNESQGGDDVATCSQDCCDCAACPAIARPOALDSTFYLQMPGSLVLIILICSVFAVVTI	300
Db	241	ARNESQGGDDVATCSQDCCDCAACPAIARPOALDSTFYLQMPGSLVLIILICSVFAVVTI	300
QY	301	LLVGRFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV	360
Db	301	LLVGRFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV	360
QY	361	IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQFGFPFRTNQVILTAENRSY	420
Db	361	IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQFGFPFRTNQVILTAENRSY	420
QY	421	RYDSLILGPNFGSILDLLELLEQLERHLQVWSPEAQRNISLQDICVAPLNPDNT	480
Db	421	RYDSLILGPNFGSILDLLELLEQLERHLQVWSPEAQRNISLQDICVAPLNPDNT	480
QY	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL	540
QY	541	SCWADYGAPVFPFLATGGYKGDYSEABALIMFTSLNNYPAGDPRLAQAKLWEAEFLM	600
Db	541	SCWADYGAPVFPFLATGGYKGDYSEABALIMFTSLNNYPAGDPRLAQAKLWEAEFLM	600
QY	601	RAFQRMAGMFQVTFMAERSLEINRTTAEDLPIFATSYVIVFLYISLALGSYSWSRV	660

Db	601	RAFQRMAGMFQVTFMAERSLEINRTTAEDLPIFATSYVIVFLYISLALGSYSWSRV	660
QY	661	MVDSKATILGLGGVAVVLGAVMAANGPFYSILGIRSLVILQVVPFLVLSVGADNIFIFVLE	720
Db	661	MVDSKATILGLGGVAVVLGAVMAANGPFYSILGIRSLVILQVVPFLVLSVGADNIFIFVLE	720
QY	721	YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFLICALTPMPAVRTPALTSGLAV	780
Db	721	YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFLICALTPMPAVRTPALTSGLAV	780
QY	781	ILDFLLQMSAFVALLSLDSKQEAASRLDVCVCKPQELPPPGQEGGLLGFQKAYAPFL	840
Db	781	ILDFLLQMSAFVALLSLDSKQEAASRLDVCVCKPQELPPPGQEGGLLGFQKAYAPFL	840
QY	841	LHWITRGVVLFLALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVVLFLALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
QY	901	YFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIYATEPPEQSYLAIPASSWVDDFIDW	960
Db	901	YFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIYATEPPEQSYLAIPASSWVDDFIDW	960
QY	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLWFLNDRP	1020
QY	1021	NIKCPKGLAAYSTVNLTSQGVLT-----ASRFMAVH	1053
Db	1021	NIKCPKGLAAYSTVNLTSQGVLT-----ASRFMAVH	1053
QY	1054	KPLKNSQDYTBALRAARELANITADLRKVPCTDPAFVFPVTITNVFEQVLTILPEGL	1113
Db	1054	KPLKNSQDYTBALRAARELANITADLRKVPCTDPAFVFPVTITNVFEQVLTILPEGL	1113
QY	1081	FMLSCLVPTFAVSCLLGLDLRSLNLLSIWMLVDTVGFMAWDSYNNAVSLINLVS	1173
Db	1081	FMLSCLVPTFAVSCLLGLDLRSLNLLSIWMLVDTVGFMAWDSYNNAVSLINLVS	1173
QY	1141	AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAFAGVAMTNLPGILVLGLAKAQ	1233
Db	1141	AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAFAGVAMTNLPGILVLGLAKAQ	1233
QY	1201	LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPVDPNPALEQKRAEAAVAVVASCP	1293
Db	1201	LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPVDPNPALEQKRAEAAVAVVASCP	1293
QY	1234	LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPVDPNPALEQKRAEAAVAVVASCP	1332
Db	1234	LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPVDPNPALEQKRAEAAVAVVASCP	1332
QY	1294	NHPSRVSTADNIYVNHSPGSIKGAGAI SNFLPNNGRQF	1359
Db	1294	NHPSRVSTADNIYVNHSPGSIKGAGAI SNFLPNNGRQF	1359

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLASISYKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLASISYKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
QY 181 ATAVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATAVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQDDVATCSQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESQDDVATCSQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFPFRTNOVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFPFRTNOVILTAPNRSY 420
QY 421 RYDSSLGPNKSGIILDLLELLELQRLRLHQLWSPEAQRNLSQDICVAPLNPDNT 480
DB 421 RYDSSLGPNKSGIILDLLELLELQRLRLHQLWSPEAQRNLSQDICVAPLNPDNT 480
QY 481 SLVDCINSILQYFQNNRTLLLTANQTLWGQTSQVDKQHLFCANAPLTKQGTALAL 540
DB 481 SLVDCINSILQYFQNNRTLLLTANQTLWGQTSQVDKQHLFCANAPLTKQGTALAL 540
QY 541 SCHADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLWEEAPLEB 600
DB 541 SCHADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLWEEAPLEB 600
QY 601 RAFORMAGMPOVTFMAERSLDEINRTTAEDLPATSYIVIFLYISALGYSWSRV 660
DB 601 RAFORMAGMPOVTFMAERSLDEINRTTAEDLPATSYIVIFLYISALGYSWSRV 660
QY 661 WYDSKATLGLGVAVVLGAVMAAMGPPSYLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
DB 661 WYDSKATLGLGVAVVLGAVMAAMGPPSYLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEALICPFLGALTMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEALICPFLGALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKQELPPQGGEGLLGFPQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKQELPPQGGEGLLGFPQKAYAPFL 840
QY 841 LHMWITRGVLLALFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVAPV 900
DB 841 LHMWITRGVLLALFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVAPV 900
QY 901 YFVITLGYNFSBAGNNAICSSAGCNPFSTQKIYVATEFPQSYLAIPASSWVDDFDIW 960
DB 901 YFVITLGYNFSBAGNNAICSSAGCNPFSTQKIYVATEFPQSYLAIPASSWVDDFDIW 960
QY 961 LTFPSSCCRLYISGPNKDFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLFWFLNDRP 1020
DB 961 LTFPSSCCRLYISGPNKDFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLFWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053

QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEFVFPYTTITNVFYEQYLTIPEGL 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEFVFPYTTITNVFYEQYLTIPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLLNLSIVMILVDTVGFMAWMDISYNAVSLNLVS 1173
DB 1141 FMLSCLVPTFAVSCLLGLDLRLSGLLNLSIVMILVDTVGFMAWMDISYNAVSLNLVS 1200
QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIPFFRLNLITLGLLHGLVFLPVILSYVYGPVNPALALEQKEAEBAVAAVMVASCP 1293
DB 1261 LIQIPFFRLNLITLGLLHGLVFLPVILSYVYGPVNPALALEQKEAEBAVAAVMVASCP 1320
QY 1294 NHPSRVSTADNTYVNHSPFGSIKAGAIISNPLPNNGRQF 1332
DB 1321 NHPSRVSTADNTYVNHSPFGSIKAGAIISNPLPNNGRQF 1359

RESULT 19

PCT-US05-27579-22
; Sequence 22, Application PC/TUS0527579
; GENERAL INFORMATION:
; APPLICANT: Mount Sinai School of Medicine
; APPLICANT: Ioannou, Yiannis
; APPLICANT: Davies, Joanna P.
; TITLE OF INVENTION: NPC1L1 AND NPC1L1 INHIBITORS AND METHODS OF USE THEREOF
; FILE REFERENCE: 2201581-WOO
; CURRENT APPLICATION NUMBER: PCT/US05/27579
; CURRENT FILING DATE: 2005-08-12
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-27579-22

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLASISYKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLASISYKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
QY 181 ATAVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATAVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQDDVATCSQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESQDDVATCSQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHFGFPFRTNOVILTAPNRSY 420

```
Db 361 IPVVALAAGLVFTELTTPDVELMSAPNSQARSEKAFHQHFGPPRRTNQVILTAPNRSSY 420
QY 421 RYDSLLLPKPNPGLDLDLLELELQERLRHLQVWSPBAQRNLSLQDICYAPLNPNT 480
Db 421 RYDSLLLPKPNPGLDLDLLELELQERLRHLQVWSPBAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKQGTALAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEM 600
Db 541 SCWADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAFORMMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISLGSYSWSRV 660
Db 601 RAFORMMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISLGSYSWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFESYLGIRSSLVILQVPRVLVSGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFESYLGIRSSLVILQVPRVLVSGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLQSKQAEASRLDVCCCKPQELPPPGQGGGLLGGFPQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLQSKQAEASRLDVCCCKPQELPPPGQGGGLLGGFPQKAYAPPL 840
QY 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
Db 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNFFSTQKI QVATEFPFQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTLGYNFSSEAGNNAICSSAGCNFFSTQKI QVATEFPFQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYL PWF LNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYL PWF LNDRP 1020
QY 1021 NIKCPKGGAAVSTSNLTSQGVLT-----ASRPMAYH 1053
Db 1021 NIKCPKGGAAVSTSNLTSQGVLT-----ASRPMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEFVFPYTIITNVFYEQLTILPEGL 1113
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEFVFPYTIITNVFYEQLTILPEGL 1113
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLLNLLSIVMLIVDTVGMALWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRLSGLLNLLSIVMLIVDTVGMALWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVPAGVAMTNLPGIILVLGAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVPAGVAMTNLPGIILVLGAKAQ 1260
QY 1234 LIQIFPFRNLITLTLGLHLGLVFLPVLVSYGPDVNPALAEQKAEAEAAVAVVASC 1293
Db 1261 LIQIFPFRNLITLTLGLHLGLVFLPVLVSYGPDVNPALAEQKAEAEAAVAVVASC 1320
QY 1294 NHPSRVSTADNIYVNHSPFGSIIKAGAI SNF L PNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIIKAGAI SNF L PNNGRQF 1359
```

RESULT 20

```
US-10-258-899A-1830
; Sequence 1830, Application US/10258899A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Cao, Yicheng
; APPLICANT: Ma, Yunquing
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Xue, Aidong
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Goodrich, Ryle
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 787CIP2-2G/US
; CURRENT APPLICATION NUMBER: US/10/258,899A
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: PCT/US01/04098
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/774,434
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; TYPE: PRT
; LENGTH: 1359
; ORGANISM: Homo sapiens
US-10-258-899A-1830
```

Query Match 99.5%; Score 6872.5; DB 32; Length 1359;

Best Local Similarity 97.9%; Pred No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

```
QY 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLQKICPRLYTGNTQAACSAKQLVSLASLSTIKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPRLYTGNTQAACSAKQLVSLASLSTIKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAVVAEYAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAVVAEYAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATAVGTWCGVYGSALCNARWLNFGQDTNGCLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db 181 ATAVGTWCGVYGSALCNARWLNFGQDTNGCLAPLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNSQGGDVATCSCQDCNAACSPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGGDVATCSCQDCNAACSPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
```


Db 481 SLVDCINSLQYFONNETLTLTANQTLMGTSQVDMKDHFLYCANAPLTFKGTALAL 540
QY 541 SCMDYGAAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAOKLWEEAFLEEM 600
Db 541 SCMDYGAAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAOKLWEEAFLEEM 600
QY 601 RAQORMAGMFOVTPFAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSWSRV 660
Db 601 RAQORMAGMFOVTPFAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSWSRV 660
QY 661 MYDSKATLGLGVAVVLGAVMAAMGFFSVLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFFSVLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQGGELLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQGGELLLGFFQKAYAPFL 840
QY 841 LHMWTRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHMWTRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTLTGYNFSSEAGMNAICSSAGCNFSFTQKI QYATEFPQSYLAIPASSWVDDFDW 960
Db 901 YFVTLTGYNFSSEAGMNAICSSAGCNFSFTQKI QYATEFPQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQFHKLDPFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQFHKLDPFLNDRP 1020
QY 1021 NIKCPKGGGLAAYSTVNLTSDDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGGGLAAYSTVNLTSDDQVLTVALSPRLEYSCTISAHCNLYLLDSASRFMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANIADLRKVPGTDPAFEPVFTYITNVFVEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANIADLRKVPGTDPAFEPVFTYITNVFVEQVLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLSRLNLSITVMTLVDVTFGMALWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLSRLNLSITVMTLVDVTFGMALWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIPFFRNLLITLGLHLGLVFLPVILSYVGPDPVNPALAEQKAEBAVAAMVASCP 1293
Db 1261 LIQIPFFRNLLITLGLHLGLVFLPVILSYVGPDPVNPALAEQKAEBAVAAMVASCP 1320
QY 1294 NHPRSVSTADNIVNHSFGSITKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPRSVSTADNIVNHSFGSITKGAGAI SNFLPNNGRQF 1359

RESULT 22

US-10-621-758A-44
; Sequence 44, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-44

Query Match 99.5%; Score 6872.5; DB 36; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHILILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSPNQSLFINVTVAQOLGAGQLPVAVYEAFOHSFAEQSYDCSRVRPAA 180
Db 121 VNLCHNTCSPNQSLFINVTVAQOLGAGQLPVAVYEAFOHSFAEQSYDCSRVRPAA 180
QY 181 ATLAVGTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITTFHLLBFGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITTFHLLBFGQAVGSGIQPLNEGV 240
QY 241 ARCNEQDQDDVATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQDQDDVATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSNQVDPKKGTSLSDKLSFTHTLLGQFQGMGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSNQVDPKKGTSLSDKLSFTHTLLGQFQGMGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPDELWSAPNSQASEKAFHQHFGFPFRTNQVILTAENRSSY 420
Db 361 IPVVALAAGLVFTLTTPDELWSAPNSQASEKAFHQHFGFPFRTNQVILTAENRSSY 420
QY 421 RYDLSLLGPKNFSGLDLDLLELELERLHQVWSPQAQRNLSQDI CYPALNPONT 480
Db 421 RYDLSLLGPKNFSGLDLDLLELELERLHQVWSPQAQRNLSQDI CYPALNPONT 480
QY 481 SLYDCCINSLQYFONNETLTLTANQTLMGTSQVDMKDHFLYCANAPLTFKGTALAL 540
Db 481 SLYDCCINSLQYFONNETLTLTANQTLMGTSQVDMKDHFLYCANAPLTFKGTALAL 540
QY 541 SCMDYGAAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAOKLWEEAFLEEM 600
Db 541 SCMDYGAAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAOKLWEEAFLEEM 600
QY 601 RAQORMAGMFOVTPFAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSWSRV 660
Db 601 RAQORMAGMFOVTPFAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSWSRV 660
QY 661 MYDSKATLGLGVAVVLGAVMAAMGFFSVLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFFSVLGRSSSLVILQVVPFLVSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQGGELLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQGGELLLGFFQKAYAPFL 840
QY 841 LHMWTRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHMWTRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

QY 901 YFVTTTLYGNFSSAGMNAICSSAGCNFFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDW 960
 Db 901 YFVTTTLYGNFSSAGMNAICSSAGCNFFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDW 960
 QY 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMCSITMGSVRPSVEQFHKYLIPWFLNDRP 1020
 Db 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMCSITMGSVRPSVEQFHKYLIPWFLNDRP 1020
 QY 1021 NIKCPKGLAAVSTSVNLTSDGVL-----ASRFMAVH 1053
 Db 1021 NIKCPKGLAAVSTSVNLTSDGVLDTVAILSPRLEYSGTISAHCNLYLLDSASRFMAVH 1080
 QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1113
 Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1140
 QY 1114 FMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILVDTVGPMAWDISYNAVSLINLVS 1173
 Db 1141 FMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILVDTVGPMAWGISYNAVSLINLVS 1200
 QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
 Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
 QY 1234 LIQIPFFRLNLLITLGLLHGLVFLPVLISYVGPVNPALALEQKRAEBAVAAMVVASCP 1293
 Db 1261 LIQIPFFRLNLLITLGLLHGLVFLPVLISYVGPVNPALALEQKRAEBAVAAMVVASCP 1320
 QY 1294 NHPRSVSTADNIVNHSFGSFKGAGAI SNFLPNNGRQF 1332
 Db 1321 NHPRSVSTADNIVNHSFGSFKGAGAI SNFLPNNGRQF 1359
 RESULT 23
 US-10-646-301A-44
 ; Sequence 44, Application US/10646301A
 ; GENERAL INFORMATION:
 ; APPLICANT: Altmann, Scott W
 ; APPLICANT: Wang, Luquan
 ; APPLICANT: Graziano, Michael
 ; APPLICANT: Murgolo, Nick
 ; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
 ; FILE REFERENCE: JB01603-K1-US
 ; CURRENT APPLICATION NUMBER: US/10/646,301A
 ; CURRENT FILING DATE: 2003-08-22
 ; PRIOR APPLICATION NUMBER: 60/397,442
 ; PRIOR FILING DATE: 2002-07-19
 ; PRIOR APPLICATION NUMBER: 10/621,758
 ; PRIOR FILING DATE: 2003-07-17
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 44
 ; LENGTH: 1359
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-646-301A-44
 Query Match 99.5%; Score 6872.5; DB 36; Length 1359;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
 QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
 Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
 QY 61 TPARKITGDHLLILLOKICRLPYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
 Db 61 TPARKITGDHLLILLOKICRLPYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
 QY 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
 Db 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180

QY 181 ATLAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLIDITFHLLEPCQAVSGSIQPLNEGV 240
 Db 181 ATLAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLIDITFHLLEPCQAVSGSIQPLNEGV 240
 QY 241 ARCNSQGGDDVATCSQDCAACPAIARPOALDSTFYLQMPGSLVLIILCSVFAVVTI 300
 Db 241 ARCNSQGGDDVATCSQDCAACPAIARPOALDSTFYLQMPGSLVLIILCSVFAVVTI 300
 QY 301 LLVGFVRVAPARDKSMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
 Db 301 LLVGFVRVAPARDKSMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
 QY 361 IPVVALAAGLVTELTTPDELWSAPNSQARSEKAFHDQHPFPFRTNVOVILTAPNRSSY 420
 Db 361 IPVVALAAGLVTELTTPDELWSAPNSQARSEKAFHDQHPFPFRTNVOVILTAPNRSSY 420
 QY 421 RYDSLGLGPNFSGILDLDDLLELLELQERLRLHQLVMSPEAQRNLSQDICVAPLNDPT 480
 Db 421 RYDSLGLGPNFSGILDLDDLLELLELQERLRLHQLVMSPEAQRNLSQDICVAPLNDPT 480
 QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
 Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
 QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLEAEFLSEM 600
 Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLEAEFLSEM 600
 QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSIVIVFLYISALGSSWSRV 660
 Db 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFATSIVIVFLYISALGSSWSRV 660
 QY 661 MYDSKATLGLGGVAVVVLGAVMAAGFPFSLGRSSILVQLQVVPFLVLSGADNIFLVL 720
 Db 661 MYDSKATLGLGGVAVVVLGAVMAAGFPFSLGRSSILVQLQVVPFLVLSGADNIFLVL 720
 QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
 Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
 QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPRQGGEGLLIGFQKAYAPFL 840
 Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPRQGGEGLLIGFQKAYAPFL 840
 QY 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEGAPV 900
 Db 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEGAPV 900
 QY 901 YFVTTTLYGNFSSAGMNAICSSAGCNFFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDW 960
 Db 901 YFVTTTLYGNFSSAGMNAICSSAGCNFFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDW 960
 QY 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMCSITMGSVRPSVEQFHKYLIPWFLNDRP 1020
 Db 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMCSITMGSVRPSVEQFHKYLIPWFLNDRP 1020
 QY 1021 NIKCPKGLAAVSTSVNLTSDGVL-----ASRFMAVH 1053
 Db 1021 NIKCPKGLAAVSTSVNLTSDGVLDTVAILSPRLEYSGTISAHCNLYLLDSASRFMAVH 1080
 QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1113
 Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1140
 QY 1114 FMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILVDTVGPMAWDISYNAVSLINLVS 1173
 Db 1141 FMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILVDTVGPMAWGISYNAVSLINLVS 1200
 QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
 Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
 QY 1234 LIQIPFFRLNLLITLGLLHGLVFLPVLISYVGPVNPALALEQKRAEBAVAAMVVASCP 1293

Db 1261 LIQIFFFRLNLLITLLGLHLGLVFLPVILSYGPDVNPALALEQKRAEAAVAVMVASCP 1320
Qy 1284 NHPSRVSTADNIYVNHSPFGSIKGAAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKGAAGAI SNFLPNNGRQF 1359

RESULT 24
US-10-663-208A-44
; Sequence 44, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-208A-44

Query Match 99.5%; Score 6872.5; DB 36; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLWALLRLLAQSPYTTIHQPGYCAFYDECGKRPGLSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSPYTTIHQPGYCAFYDECGKRPGLSGSLMTLSNVCSLSN 60
Qy 61 TPARKITGDHLLILLOKICRPLYTGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICRPLYTGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCNEGQDDVATCSQDCACAPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEGQDDVATCSQDCACAPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGFOGWTGWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGFOGWTGWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGPFRTNQVILITAPNRSY 420
Qy 421 RYDLSLLGPNKSGIILDLLELLEQLERHLQVWSPQARNISLQDICVAPLNPDNT 480
Db 421 RYDLSLLGPNKSGIILDLLELLEQLERHLQVWSPQARNISLQDICVAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540

Qy 541 SCMDYGAIPVPPFLAIGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAFLEBEM 600
Db 541 SCMDYGAIPVPPFLAIGYKGYSEAEALIMTSLNNYPAGDPRLAQAKLWEAFLEBEM 600
Qy 601 RAFQRMAGMFQVTTFAERSLEDBINRTTAEADLPFATSYIVIFLYISLALGSSYSWSRV 660
Db 601 RAFQRMAGMFQVTTFAERSLEDBINRTTAEADLPFATSYIVIFLYISLALGSSYSWSRV 660
Qy 661 MYDSKATILGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGDADNIFVLE 720
Db 661 MYDSKATILGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGDADNIFVLE 720
Qy 721 YORLPRRPGEPREVIHGRALGRVAPSMCLLSLSAICFFLALTPMPAVRFAITSLGLAV 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMCLLSLSAICFFLALTPMPAVRFAITSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVXPQELPPPPQGGEGLLGFPQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVXPQELPPPPQGGEGLLGFPQKAYAPFL 840
Qy 841 LHWITRGVVLFLALFGVSLYSKCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLALFGVSLYSKCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVITLGYNPSSEAGMNAICSSAGCNPFSFTQKIYATEFPEQSYLAIPASSWVDDFDW 960
Db 901 YFVITLGYNPSSEAGMNAICSSAGCNPFSFTQKIYATEFPEQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLWPFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKYLWPFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTVNLTSQGVL-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAYSTVNLTSQGVLDTVAILSPRLEYSGTISAHCNLYLLDSASRPMAYH 1080
Qy 1054 KPLKNSQDYTBALRAARELAANIADLRKVPCTDPAPVFPVYITNVFVEQYLTILPEGL 1113
Db 1054 KPLKNSQDYTBALRAARELAANIADLRKVPCTDPAPVFPVYITNVFVEQYLTILPEGL 1113
Qy 1114 FMLSCLVPTFAVSCLLGLRLSGLNLLSIVMLVDTVGPALMDISYNAVSLINLVS 1173
Db 1114 FMLSCLVPTFAVSCLLGLRLSGLNLLSIVMLVDTVGPALMDISYNAVSLINLVS 1173
Qy 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVGLAKAQ 1233
Db 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVGLAKAQ 1233
Qy 1234 LIQIPFFRLNLLITLLGLHLGLVFLPVILSYGPDVNPALALEQKRAEAAVAVMVASCP 1293
Db 1234 LIQIPFFRLNLLITLLGLHLGLVFLPVILSYGPDVNPALALEQKRAEAAVAVMVASCP 1293
Qy 1294 NHPSRVSTADNIYVNHSPFGSIKGAAGAI SNFLPNNGRQF 1332
Db 1294 NHPSRVSTADNIYVNHSPFGSIKGAAGAI SNFLPNNGRQF 1359

RESULT 25
US-10-736-769-44
; Sequence 44, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758

;
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-736-769-44

Query Match 99.5%; Score 6872.5; DB 37; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVYAEFYQHSFABQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVYAEFYQHSFABQSYDSCSRVRPAA 180

QY 181 ATLAUVTGCVYSALCNAQRLNFGQDTGNGLAFLDITPHLLPEQAVGSGIQLNEGV 240
Db 181 ATLAUVTGCVYSALCNAQRLNFGQDTGNGLAFLDITPHLLPEQAVGSGIQLNEGV 240

QY 241 ARCNEQSGDDVATCSQDCAASCPAARQALDSTFYLQMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNEQSGDDVATCSQDCAASCPAARQALDSTFYLQMPGSLVLIILCSVPAVVTI 300

QY 301 LVGFRVAPARDKSKWMDPKGTSLSDLKSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LVGFRVAPARDKSKWMDPKGTSLSDLKSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTELTDPVELMSAPNSQARSEKAFHDQHPGFFRTNQVILTAPEKRSY 420
Db 361 IPVVALAAGLVFTELTDPVELMSAPNSQARSEKAFHDQHPGFFRTNQVILTAPEKRSY 420

QY 421 RYDSSLGPKNPSGIILDLLELLELQERLRLHQLVMSPEAQRNISLQDICVAPLNPNT 480
Db 421 RYDSSLGPKNPSGIILDLLELLELQERLRLHQLVMSPEAQRNISLQDICVAPLNPNT 480

QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTKDGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTKDGTALAL 540

QY 541 SCHADYGAPVFPFLAIGYKGDYSEAEALIMTFPSLNNTYAGDPRLAQAQKWEAPLEEM 600
Db 541 SCHADYGAPVFPFLAIGYKGDYSEAEALIMTFPSLNNTYAGDPRLAQAQKWEAPLEEM 600

QY 601 RAFORMAGMFWTTFATERSLDEINRTTAEDLPATSYIVIFLYISIALGYSWSRV 660
Db 601 RAFORMAGMFWTTFATERSLDEINRTTAEDLPATSYIVIFLYISIALGYSWSRV 660

QY 661 MVDSKATLGLGGVAVVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIFVLE 720
Db 661 MVDSKATLGLGGVAVVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIFVLE 720

QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEATCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEATCFPLGALTMPAVRTFALTSLGLAV 780

QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDYCCCKVQOELPPGQEGILLGPFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDYCCCKVQOELPPGQEGILLGPFQKAYAPFL 840

QY 841 LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
Db 841 LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900

QY 901 YFVTTILGYNPSSEAGNNAICSSAGCNPFSTQKIQYATPEPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNPSSEAGNNAICSSAGCNPFSTQKIQYATPEPQSYLAIPASSWVDDFIDW 960

QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCILKNCMSITMGSVRPSVQFHKYLPWFNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCILKNCMSITMGSVRPSVQFHKYLPWFNDRP 1020

QY 1021 NIKCPKGLAAYSTSVNLTSDDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDDQVL-----ASRFMAYH 1053

QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFEVPPYTTITNVFVQYLTILPEGL 1113
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFEVPPYTTITNVFVQYLTILPEGL 1113

QY 1114 FMLSCLVPTFAVSCILLGLDLRSGLLNLISIVMILVDTVGFMAWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCILLGLDLRSGLLNLISIVMILVDTVGFMAWDISYNAVSLINLVS 1200

QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260

QY 1234 LIQIPFRLNLTITLGLLHGLVFLPVILSYGPDVNPALALBOKRAEBAVAAVMVASCP 1293
Db 1261 LIQIPFRLNLTITLGLLHGLVFLPVILSYGPDVNPALALBOKRAEBAVAAVMVASCP 1320

QY 1294 NHPSRVSTADNIYVNHSPFGSIKAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAI SNFLPNNGRQF 1359

RESULT 26

US-10-750-386-44
; Sequence 44, Application US/10750386
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; APPLICANT: Chapman, Kevin
; APPLICANT: Goulet, Mark
; APPLICANT: Ujainwalla, Peroze
; APPLICANT: Altmann, Scott W
; APPLICANT: Davis, Chip
; APPLICANT: Bull, Herb
; APPLICANT: Thornberry, Nancy A
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONIST
; FILE REFERENCE: A36104 074669 0103
; CURRENT APPLICATION NUMBER: US/10750,386
; CURRENT FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-750-386-44

Query Match 99.5%; Score 6872.5; DB 37; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
DB 121 VNLHCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAGVTMCGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLPEQAVGSGIQPLNEGV 240
DB 181 ATLAGVTMCGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNSQGGDDVATCSCODCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNSQGGDDVATCSCODCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLWSPEAQRNLSLODICVAPLNPNT 480
DB 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLWSPEAQRNLSLODICVAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDWKDHELYCANAPLTFKGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDWKDHELYCANAPLTFKGTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEAEFLBEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEAEFLBEM 600
QY 601 RAFORMMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIIVIFYLSIALGYSYSWSRV 660
DB 601 RAFORMMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIIVIFYLSIALGYSYSWSRV 660
QY 661 MYDSKATILGGLGVAVVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIIFFVLE 720
DB 661 MYDSKATILGGLGVAVVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIIFFVLE 720
QY 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPQCGEGILLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPQCGEGILLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSEAGMNAICSSAGCANNFSTQKIQVATEFFPQSYLAIPASSWVDIDF 960
DB 901 YFVTTILGYNFSSEAGMNAICSSAGCANNFSTQKIQVATEFFPQSYLAIPASSWVDIDF 960
QY 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCIKNCSITMGSVRPSVEQFKHYLPWFANDRP 1020
DB 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCIKNCSITMGSVRPSVEQFKHYLPWFANDRP 1020
QY 1021 NTKCPKGGGLAAVSTSNVLSIDGQVL-----ASRPMAYH 1053
DB 1021 NTKCPKGGGLAAVSTSNVLSIDGQVL-----ASRPMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNVFYEQVLTILPEGL 1113
DB 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVPTITNVFYEQVLTILPEGL 1113
QY 1114 FMLSCLVPTFAVSCILLGLDLSGLNLLSIIMVILVDTVGFMAWDISYNAVSLINLYS 1173
DB 1141 FMLSCLVPTFAVSCILLGLDLSGLNLLSIIMVILVDTVGFMAWDISYNAVSLINLYS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1233

DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIFFFFRLNLITLLGLLHGLVFLPVLTLSVVGDDVNPALALEOKRABEAAVAVNVASCP 1293
DB 1261 LIQIFFFFRLNLITLLGLLHGLVFLPVLTLSVVGDDVNPALALEOKRABEAAVAVNVASCP 1320
QY 1294 NHPRSVSTADNIYVNHSPESIGKAGAGAISNPLPNNGRQF 1332
DB 1321 NHPRSVSTADNIYVNHSPESIGKAGAGAISNPLPNNGRQF 1359

RESULT 27

US-11-218-141-1830
; Sequence 1830. Application US/11218141
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/11/218,141
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-218-141-1830

Query Match 99.5%; Score 6872.5; DB 42; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLLOKICPRLVTGNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLLOKICPRLVTGNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
DB 121 VNLHCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAGVTMCGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLPEQAVGSGIQPLNEGV 240
DB 181 ATLAGVTMCGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNSQGGDDVATCSCODCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNSQGGDDVATCSCODCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWASWPLTILVLSV 360

```
Db 301 LLVGRVAPARDKSKWDPKKGTSLSKUSFSFTHLLGQFPQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTTELTPDPVELMSAPNSQARSEKAFHDQHFQGFPRFTNQVILTAPNRSSV 420
Db 361 IPVVALAAGLVFTTELTPDPVELMSAPNSQARSEKAFHDQHFQGFPRFTNQVILTAPNRSSV 420
Qy 421 RYDSLLGPKNFGSGLDLDLLELLELQERLRLHQLWSPQAQNSISLQDICYAPLNPDMT 480
Db 421 RYDSLLGPKNFGSGLDLDLLELLELQERLRLHQLWSPQAQNSISLQDICYAPLNPDMT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Qy 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
Db 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
Qy 601 RAPQRMAGMFOVTFMAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660
Db 601 RAPQRMAGMFOVTFMAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAMGPFYSYLGRSSVLVILQVVPFLVLSYGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAMGPFYSYLGRSSVLVILQVVPFLVLSYGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSGSLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSGSLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQBSRLDVCCVKPQELPPPGQEGELGFGFKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQBSRLDVCCVKPQELPPPGQEGELGFGFKAYAPFL 840
Qy 841 LHWITRGVLLPLALPGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVEGAPV 900
Db 841 LHWITRGVLLPLALPGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVEGAPV 900
Qy 901 YFVTTLYGNFSSBAGNNAICSSAGCNPFSTOKIOYATPFPQSYLAIPASSWVDDFDIM 960
Db 901 YFVTTLYGNFSSBAGNNAICSSAGCNPFSTOKIOYATPFPQSYLAIPASSWVDDFDIM 960
Qy 961 LTPSSCCRLYISGPNKDFECPSTVNSLNCNKMSITMGSVRPSVEQFHKYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFECPSTVNSLNCNKMSITMGSVRPSVEQFHKYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLT-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLTVAILSPRLEYSGTISAHCNLYLLDSASRPMAYH 1080
Qy 1054 KPLKNSODYTEALRAARELAANITADLRKVGCTDPAPFPYPTITNVFYEQYITLPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAPFPYPTITNVFYEQYITLPEGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDVTGFMALWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDVTGFMALWDISYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIPFRLNLLITLGLHGLVFLPVILSYGPDVNPALAEOKRAEBAVAAMVASCPC 1293
Db 1261 LIQIPFRLNLLITLGLHGLVFLPVILSYGPDVNPALAEOKRAEBAVAAMVASCPC 1320
Qy 1294 NHPSRVSTADNTVYVHNSFEGSIKGAGAI SNFLPNNGROF 1332
Db 1321 NHPSRVSTADNTVYVHNSFEGSIKGAGAI SNFLPNNGROF 1359
```

```
US-10-170-205E-21686
; Sequence 21686, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 21686
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-21686
```

Query Match 99.4%; Score 6868.5; DB 31; Length 1359;
Best Local Similarity 97.8%; Pred. No. 0;

```
Matches 1329; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLWALLRLAQAQSEPYTTIHOPGYCAFYDECGKNPGLSGSLMTLSNVCLSN 60
Db 1 MAEAGLRGWLWALLRLAQAQSEPYTTIHOPGYCAFYDECGKNPGLSGSLMTLSNVCLSN 60
Qy 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNQSLFINVTRVAQAGAGQLPAVAYEAFYQHSFABQSYDSCSRVRPAA 180
Db 121 VNLHCHTCSNQSLFINVTRVAQAGAGQLPAVAYEAFYQHSFABQSYDSCSRVRPAA 180
Qy 181 ATLAVGTMCVGVGSALCNAQRWLNFGDGTNGNLAPLDITFHLLEPQAVGSGIQLNEGV 240
Db 181 ATLAVGTMCVGVGSALCNAQRWLNFGDGTNGNLAPLDITFHLLEPQAVGSGIQLNEGV 240
Qy 241 ARCNSQGDVATCSQDCAACPAIARPOALDSTFYLQMPGSLVLIILCSVPFVAVTI 300
Db 241 ARCNSQGDVATCSQDCAACPAIARPOALDSTFYLQMPGSLVLIILCSVPFVAVTI 300
Qy 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTTELTPDPVELMSAPNSQARSEKAFHDQHFQGFPRFTNQVILTAPNRSSV 420
Db 361 IPVVALAAGLVFTTELTPDPVELMSAPNSQARSEKAFHDQHFQGFPRFTNQVILTAPNRSSV 420
Qy 421 RYDSLLGPKNFGSGLDLDLLELLELQERLRLHQLWSPQAQNSISLQDICYAPLNPDMT 480
Db 421 RYDSLLGPKNFGSGLDLDLLELLELQERLRLHQLWSPQAQNSISLQDICYAPLNPDMT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Qy 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
Db 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
Qy 601 RAPQRMAGMFOVTFMAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660
Db 601 RAPQRMAGMFOVTFMAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAMGPFYSYLGRSSVLVILQVVPFLVLSYGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAMGPFYSYLGRSSVLVILQVVPFLVLSYGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSGSLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSGSLAV 780
```

```
QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKPQELPPPPQGGELLLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKPQELPPPPQGGELLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLGNFSSSEAGNNAICSSAGCNPFSTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGNFSSSEAGNNAICSSAGCNPFSTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNKMCSITMGSVRPSVEQFHKYLPPWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNKMCSITMGSVRPSVEQFHKYLPPWFLNDRP 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053
DB 1021 NIKCPKGGLAAYSTSVNLTSDGQVLTVAILSPRLEYSOTISAHCNLYLLDSTSRFMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTIITNVFEQYVLTILPEGL 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTIITNVFEQYVLTILPEGL 1140
QY 1114 FMLSICLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGFMAWLDISYNAVSLINLVS 1173
DB 1141 FMLSICLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGFMAWLDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1260
QY 1234 LIQIFFRNLNLTILGLLHGLVFLPVLISYVGPDPVPALAEQKRAEBAVAANVASC 1293
DB 1261 LIQIFFRNLNLTILGLLHGLVFLPVLISYVGPDPVPALAEQKRAEBAVAANVASC 1320
QY 1294 NHPRSVSTADNLYVNHSPGSIKAGAINFLPNNGRQF 1332
DB 1321 NHPRSVSTADNLYVNHSPGSIKAGAINFLPNNGRQF 1359

RESULT 29
US-60-453-050-8956
; Sequence 8956, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8956

Query Match 99.4%; Score 6868.5; DB 48; Length 1359;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLBASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLBASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSDSCSRVRPAA 180
```

```
DB 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSDSCSRVRPAA 180
QY 181 ATLAGVTMCYGSALCNAQRWLNFGQDTGNGLAPLDTIHELLPQGAQVSGIQLPNEG 240
DB 181 ATLAGVTMCYGSALCNAQRWLNFGQDTGNGLAPLDTIHELLPQGAQVSGIQLPNEG 240
QY 241 ARCNESQDDVATCSQDCAASCPAIARPQALDSTFYLGQMPGSLVLIILCSVFVV 300
DB 241 ARCNESQDDVATCSQDCAASCPAIARPQALDSTFYLGQMPGSLVLIILCSVFVV 300
QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSKLSPSTHTLLGQFFQGMGTWVASHPLTILV 360
DB 301 LLVGFVRVAPARDKSMVDPKKGTSLSKLSPSTHTLLGQFFQGMGTWVASHPLTILV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFGRPTNOVLTITAPN 420
DB 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFGRPTNOVLTITAPN 420
QY 421 RYDSLLLPKNFSGILDLLELLELQERLHLQVMSPEAQNISLQDICYAPLNPDNT 480
DB 421 RYDSLLLPKNFSGILDLLELLELQERLHLQVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKOG 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKOG 540
QY 541 SCMADYGPVFPFLAIGYKGYKDYSEBALTMTFSLNNYPAGDPRLAQAKLWEEAF 600
DB 541 SCMADYGPVFPFLAIGYKGYKDYSEBALTMTFSLNNYPAGDPRLAQAKLWEEAF 600
QY 601 RAFORMAGMFOVTFMERSLEDEINRTTADLPFATSYIVIFLYISIALGYS 660
DB 601 RAFORMAGMFOVTFMERSLEDEINRTTADLPFATSYIVIFLYISIALGYS 660
QY 661 MVDKATILGLGGVAVVLGAVMAAMGFFSYLGRSLVLQVVPFLVLSVGADNIFI 720
DB 661 MVDKATILGLGGVAVVLGAVMAAMGFFSYLGRSLVLQVVPFLVLSVGADNIFI 720
QY 721 YORLPRRRGGEBREVIHGRALGRVAPSMILCSISEAICFFLGALTMPAVR 780
DB 721 YORLPRRRGGEBREVIHGRALGRVAPSMILCSISEAICFFLGALTMPAVR 780
QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKPQELPPPPQGGELLLGFFQKAY 840
DB 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKPQELPPPPQGGELLLGFFQKAY 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLN 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHIISVGLDQELALPKDSYLLDYFLFLN 900
QY 901 YFVTTLGNFSSSEAGNNAICSSAGCNPFSTQKIQYATEFPEQSYLAIPASS 960
DB 901 YFVTTLGNFSSSEAGNNAICSSAGCNPFSTQKIQYATEFPEQSYLAIPASS 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNKMCSITMGSVRPSVEQFHKYL 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNKMCSITMGSVRPSVEQFHKYL 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053
DB 1021 NIKCPKGGLAAYSTSVNLTSDGQVLTVAILSPRLEYSOTISAHCNLYLLD 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTIITNVFEQY 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTIITNVFEQY 1140
QY 1114 FMLSICLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGFMAWLDISYNAV 1173
DB 1141 FMLSICLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGFMAWLDISYNAV 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGL 1233
DB 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGL 1233
```

Db 1201 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIFPRLMLITLLGLHGLVFLPVILSYVGDVNPALALBQKRAEBAVAAMVASC 1293
Db 1261 LIQIFPRLMLITLLGLHGLVFLPVILSYVGDVNPALALBQKRAEBAVAAMVASC 1320
QY 1294 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1359

RESULT 30
US-60-453-135-8956
; Sequence 8956, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8956

Query Match 99.4%; Score 6868.5; DB 48; Length 1359;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

QY 1 MAEAGLUGMLLWALLRLAQSEPYTHOPGYCAFYDECGKPELSGLMTLSNVSLSN 60
Db 1 MAEAGLUGMLLWALLRLAQSEPYTHOPGYCAFYDECGKPELSGLMTLSNVSLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQLSPINTRVAQLGAGQLPAVAYEAFYQHSFABQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQLSPINTRVAQLGAGQLPAVAYEAFYQHSFABQSYDSCSRVRPAA 180
QY 181 ATLAVGTMCGVGSALCNAQRWLNFGDGTGNGLAPLDITPHLLPQQAQVSGIQPLNEGY 240
Db 181 ATLAVGTMCGVGSALCNAQRWLNFGDGTGNGLAPLDITPHLLPQQAQVSGIQPLNEGY 240
QY 241 ARCNSQGDVATCSQDCDCAAPALARPQALDSTFYLQMPGSLVLIILICSVPAVVTI 300
Db 241 ARCNSQGDVATCSQDCDCAAPALARPQALDSTFYLQMPGSLVLIILICSVPAVVTI 300
QY 301 LLVGRFVAPARKSMVDPKKGTSLSDKLSPFTHLLGQFPQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRFVAPARKSMVDPKKGTSLSDKLSPFTHLLGQFPQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHPFRFTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHPFRFTNQVILITAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLLLELLELQERLHQLWSPEAQRNLSQDICVAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLLLELLELQERLHQLWSPEAQRNLSQDICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCHADYGAPVFFFLAIGGKGYKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCHADYGAPVFFFLAIGGKGYKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600

QY 601 RAFQRMAGMFQVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGYSYSSMRV 660
Db 601 RAFQRMAGMFQVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGYSYSSMRV 660
QY 661 MVDKATLIGGVAVVVLGAVMAAGFPYSLGRSSVLVQVPPFLVSVGADNIFIPVLE 720
Db 661 MVDKATLIGGVAVVVLGAVMAAGFPYSLGRSSVLVQVPPFLVSVGADNIFIPVLE 720
QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSSEAI CFFLGALTTPMPAVRTEALTSLAV 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSSEAI CFFLGALTTPMPAVRTEALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQESASRLDVCCCKVQELPPPGQEGILLGFFQKAVAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQESASRLDVCCCKVQELPPPGQEGILLGFFQKAVAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSEAGMNAICSSAGCANNFSTQKIQYATFEPPEOSYLAIPASSVWDIDW 960
Db 901 YFVTTILGYNFSSEAGMNAICSSAGCANNFSTQKIQYATFEPPEOSYLAIPASSVWDIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMCNCSITMGSVRPSVQFHKYLPWFNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMCNCSITMGSVRPSVQFHKYLPWFNDRP 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDQVLT-----ASRFMAVH 1053
Db 1021 NIKCPKGGLAAYSTSVNLTSDQVLTVAILSPREYSGTISAHCNLYLLDSTSRFMAVH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVGTDPAFVPPYTTITNVFEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVGTDPAFVPPYTTITNVFEQYLTILPEGL 1140
QY 1114 FMLSCLVPTFVAVSCILLGLRLSGLLNLSVMILVDTVGFMAWDISYNAVSLNLVS 1173
Db 1141 FMLSCLVPTFVAVSCILLGLRLSGLLNLSVMILVDTVGFMAWDISYNAVSLNLVS 1200
QY 1174 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIFPRLMLITLLGLHGLVFLPVILSYVGDVNPALALBQKRAEBAVAAMVASC 1293
Db 1261 LIQIFPRLMLITLLGLHGLVFLPVILSYVGDVNPALALBQKRAEBAVAAMVASC 1320
QY 1294 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1359

RESULT 31
US-60-466-412-8956
; Sequence 8956, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-8956

Query Match		99.4%;	Score 6868.5;	DB 48;	Length 1359;
Best Local Similarity		97.8%;	Pred. No. 0;		
Matches 1329;		Conservative	0;	Mismatches	3;
				Indels	27;
				Gaps	1;
Qy	1	MAEAGRGWLLWALLRLRLAQSEPYTTIHQPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN	60		
Db	1	MAEAGRGWLLWALLRLRLAQSEPYTTIHQPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN	60		
Qy	61	TPARKITGDHLILLOKICPRLTYGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLILLOKICPRLTYGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQIAGAGOLPAVVAEAFYQHSFAEQSDSCSRVPAA	180		
Db	121	VNLHCHNTCSPNQSLFINVTRVAQIAGAGOLPAVVAEAFYQHSFAEQSDSCSRVPAA	180		
Qy	181	ATLAVGTMCVGVGSGALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240		
Db	181	ATLAVGTMCVGVGSGALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240		
Qy	241	ARCNSQGDVATCSCQDCAASCPAIRPQALDSTFYLGQMPGSLVLIILCSVFAYVTI	300		
Db	241	ARCNSQGDVATCSCQDCAASCPAIRPQALDSTFYLGQMPGSLVLIILCSVFAYVTI	300		
Qy	301	LLVGRVAPARDKSKWDPKKTSLSDKLSGSTHTLLGQFQGGWGTWASWPLTLLVLSV	360		
Db	301	LLVGRVAPARDKSKWDPKKTSLSDKLSGSTHTLLGQFQGGWGTWASWPLTLLVLSV	360		
Qy	361	IPWVALLAGLVTELTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAENRSSY	420		
Db	361	IPWVALLAGLVTELTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAENRSSY	420		
Qy	421	RYDLSLLGPKNPSGILDLLLELELQRLRHQLQWSPQARNISLQDICVAPLNPDNT	480		
Db	421	RYDLSLLGPKNPSGILDLLLELELQRLRHQLQWSPQARNISLQDICVAPLNPDNT	480		
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTKDGTALAL	540		
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCANAPLTKDGTALAL	540		
Qy	541	SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM	600		
Db	541	SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM	600		
Qy	601	RAFQRMAGMFQVTFATRSLSDEINRTTAEDLPIPATSYIIVIFLYISALGYSWSRV	660		
Db	601	RAFQRMAGMFQVTFATRSLSDEINRTTAEDLPIPATSYIIVIFLYISALGYSWSRV	660		
Qy	661	MVDSKATLGLGGVAVVLGAVMAAGPFYSYLIRSSILVILQVVPFLVLSVGADNIFIFVLE	720		
Db	661	MVDSKATLGLGGVAVVLGAVMAAGPFYSYLIRSSILVILQVVPFLVLSVGADNIFIFVLE	720		
Qy	721	YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEALICFFLGALTPMPAVRTFALTSGLAV	780		
Db	721	YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEALICFFLGALTPMPAVRTFALTSGLAV	780		
Qy	781	ILDFLLQMSAFVALLSLSKQRBASRLDVCCCKVQPELPPPGQCEGLLIGFQKAYAPFL	840		
Db	781	ILDFLLQMSAFVALLSLSKQRBASRLDVCCCKVQPELPPPGQCEGLLIGFQKAYAPFL	840		
Qy	841	LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900		
Db	841	LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900		
Qy	901	YFVTILGYNFSSEAGMNAICSSAGCNFSPTQIKIYATEFPEQSYLAIPASSWVDDFDIW	960		
Db	901	YFVTILGYNFSSEAGMNAICSSAGCNFSPTQIKIYATEFPEQSYLAIPASSWVDDFDIW	960		
Qy	961	LTPSSCCRLYISGNPKDKFCPSTVNSLNCNKMSITMGSVRPSVFQPHKYLFWFLNDRP	1020		
Db	961	LTPSSCCRLYISGNPKDKFCPSTVNSLNCNKMSITMGSVRPSVFQPHKYLFWFLNDRP	1020		
Qy	1021	NIKCPKGLAAYSTSVNLTSDBGVLT-----ASRFMAYH	1053		

Db	1021	NIKCPKGLAAYSTSVNLTSDBGVLTVAIISPRLEYSGTISAHCNLYLLDSTSRFMAYH	1080		
Qy	1054	KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVPPYTTITNVFYEQYLTLPEGL	1113		
Db	1081	KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVPPYTTITNVFYEQYLTLPEGL	1140		
Qy	1114	FMLSCLVPTPAVSCLLGLDLRSGLNLLSIVMLVDTVGFMALWDISYNAVSLINLVS	1173		
Db	1141	FMLSCLVPTPAVSCLLGLDLRSGLNLLSIVMLVDTVGFMALWDISYNAVSLINLVS	1200		
Qy	1174	AVGMSVFEVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1233		
Db	1201	AVGMSVFEVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1260		
Qy	1234	LIQIPFFRLNLLITLLGLLHGLVFLPVILSVYGVDPVPALALEQKRAEBAVAAMVASCP	1293		
Db	1261	LIQIPFFRLNLLITLLGLLHGLVFLPVILSVYGVDPVPALALEQKRAEBAVAAMVASCP	1320		
Qy	1294	NHPSRVSTADNIYVNHSPFEGSIKAGAIISNPLPNNGRQF	1332		
Db	1321	NHPSRVSTADNIYVNHSPFEGSIKAGAIISNPLPNNGRQF	1359		
RESULT 32					
PCT-US01-08631-53052					
; Sequence 53052, Application PC/TUS0108631					
; GENERAL INFORMATION:					
; APPLICANT: Hyseq, Inc					
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES					
; FILE REFERENCE: 21272-049					
; CURRENT APPLICATION NUMBER: PCT/US01/08631					
; CURRENT FILING DATE: 2001-03-30					
; PRIOR FILING DATE: 09/540,217					
; PRIOR FILING DATE: 2000-03-31					
; PRIOR APPLICATION NUMBER: 09/649,167					
; PRIOR FILING DATE: 2000-08-23					
; NUMBER OF SEQ ID NOS: 60736					
; SOFTWARE: Custom					
; SEQ ID NO 53052					
; LENGTH: 1344					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
; FEATURE:					
; NAME/KEY: DOMAIN					
; LOCATION: (901)..(917)					
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,					
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87					
; NAME/KEY: misc feature					
; LOCATION: (1)..(1344)					
; OTHER INFORMATION: Xaa = X or * as defined in Table 2					
PCT-US01-08631-53052					
Query Match		94.6%;	Score 6536;	DB 1;	Length 1344;
Best Local Similarity		96.7%;	Pred. No. 0;		
Matches 1271;		Conservative	7;	Mismatches	21;
				Indels	16;
				Gaps	4;
Qy	1	MAEAGRGWLLWALLRLRLAQSEPYTTIHQPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN	60		
Db	1	MAEAGRGWLLWALLRLRLAQSEPYTTIHQPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN	60		
Qy	61	TPARKITGDHLILLOKICPRLTYGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Db	61	TPARKITGDHLILLOKICPRLTYGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120		
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQIAGAGOLPAVVAEAFYQHSFAEQSDSCSRVPAA	180		
Db	121	VNLHCHNTCSPNQSLFINVTRVAQIAGAGOLPAVVAEAFYQHSFAEQSDSCSRVPAA	180		
Qy	181	ATLAVGTMCVGVGSGALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240		
Db	181	ATLAVGTMCVGVGSGALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240		

```
QY 241 ARCNSQDDVATCSQDCAAACP--AATARPQALDSTFYLGQMPGSLVLIILILCSVFVV 298
Db 241 ARCNSQDDVATCSQDCAAACPXPAPRPSTP--RQMPGSLVLIILILCSVFVV 298
QY 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFQFGTGWASWPLTILVL 358
Db 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFQFGTGWASWPLTILVL 358
QY 359 SVIPVVALAAGLVFTTLPDVELWSAPNSQARSEKAFHDQHFQFPFRNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTLPDVELWSAPNSQARSEKAFHDQHFQFPFRNQVILTAPNRS 418
QY 419 SYRYDSLLGPKNFSGILDLILLLELLELQERLHLQVMSPEAQRNLSLQDICYAPLND 478
Db 419 SYRYDSLLGPKNFSGILDLILLLELLELQERLHLQVMSPEAQRNLSLQDICYAPLND 478
QY 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
QY 539 ALSCHMADYGAPVFPFLAIGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEBAFLE 598
Db 539 ALSCHMADYGAPVFPFLAIGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEBAFLE 598
QY 599 EMRAFQRMAGMPQVTFATERSLEDEINRTTADLPFATSYIVIFLYISLALGSYSSWS 658
Db 599 EMRAFQRMAGMPQVTFATERSLEDEINRTTADLPFATSYIVIFLYISLALGSYSSWS 658
QY 659 RWVDSKATLGLGGVAVLVGVAAMGFPFSGYLSIRSLVILQVVPFLVLSVGADNIFIV 718
Db 659 RWVDSKATLGLGGVAVLVGVAAMGFPFSGYLSIRSLVILQVVPFLVLSVGADNIFIV 718
QY 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGCALTPMPAVRTFALTSL 778
Db 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGCALTPMPAVRTFALTSL 778
QY 779 AVILDFLLQMSAFVALLSDSKRQEARLDVCCVKPQBELPPGQEGILLGFQKAYAP 838
Db 779 AVILDFLLQMSAFVALLSDSKRQEARLDVCCVKPQBELPPGQEGILLGFQKAYAP 838
QY 839 FLLHWITRGVV---LFLALFGVSLYSMCHISVGLDQELALPKDSYLLDFLFLNRYF 894
Db 839 FLLHWITRGVVVPSQLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDFLFLNRYF 898
QY 895 EVGAPVYFTTLYGNFSSAGMNAICSSAGCNFSPFTQKIQVATEPPEOSYLAI PASSSV 954
Db 895 EVGAPVYFTTLYGNFSSAGMNAICSSAGCNFSPFTQKIQVATEPPEOSYLSL PASSSV 958
QY 955 DDFIDWLTSSCCRLYISGPNKDKFCPTVNSLNCNKMWSITMGSVRPSVBFHKLPLW 1014
Db 955 DDFIDWLTSSCCRLYISGPNKDKFCPTVNSLNCNKMWSITMGSVRPSVBFHKLPLW 1018
QY 1015 FLNDRNIKCPKGLAAYSTSNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAA 1074
Db 1015 FLNDRNIKCPKGLAAYSTSNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAA 1078
QY 1075 NITADLRKVPDTPAFAVEFPYITINVFYEOYLITLPEGLFMLSCLVPTFAVSCLLGLD 1134
Db 1075 NITADLRKVPDTPAFAVEFPYITINVFYEOYLITLPEGLFMLSCLVPTFAVSCLLGLD 1138
QY 1135 LRSGLLNLLSIWMLVDTVGFMAWMDISYNVSLNINLVSAGVMSVEFVSHITRSPAISTK 1194
Db 1135 LRSGLLNLLSIWMLVDTVGFMAWMDISYNVSLNINLVSAGVMSVEFVSHITRSPAISTK 1198
QY 1195 PTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLITLGLLHG 1254
Db 1195 PTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLITLGLLHG 1258
QY 1255 LVFLPVILSYGPDVNPALALQKRAEEA-----VAAMVASCNPSPRST 1301
Db 1255 LVFLPVILSYGPDVNPALALQKRAEEGSGSHGGLAQITPPSPQLTTSMT 1313
```

RESULT 33

```
US-10-450-763-53052
; Sequence 53052, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53052
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (901)..(917)
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1344)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53052
```

Query Match 94.6%; Score 6536; DB 34; Length 1344;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

```
QY 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKPNELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKPNELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLLOKICPRLYTGNPTQACCSAKQLVSLBASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLLOKICPRLYTGNPTQACCSAKQLVSLBASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFVQHSFAEQSYDSCSRVPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFVQHSFAEQSYDSCSRVPAA 180
QY 181 ATLAVGTWCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTWCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNSQDDVATCSQDCAAACP--AATARPQALDSTFYLGQMPGSLVLIILILCSVFVV 298
Db 241 ARCNSQDDVATCSQDCAAACPXPAPRPSTP--RQMPGSLVLIILILCSVFVV 298
QY 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFQFGTGWASWPLTILVL 358
Db 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFQFGTGWASWPLTILVL 358
QY 359 SVIPVVALAAGLVFTTLPDVELWSAPNSQARSEKAFHDQHFQFPFRNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTLPDVELWSAPNSQARSEKAFHDQHFQFPFRNQVILTAPNRS 418
QY 419 SYRYDSLLGPKNFSGILDLILLLELLELQERLHLQVMSPEAQRNLSLQDICYAPLND 478
Db 419 SYRYDSLLGPKNFSGILDLILLLELLELQERLHLQVMSPEAQRNLSLQDICYAPLND 478
QY 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
QY 539 ALSCHMADYGAPVFPFLAIGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEBAFLE 598
```



```

Db      539  ALSCWADYGPVFPFLALGGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFL 598
Qy      599  ENRAFQRMAGMFOVTFTAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWS 658
Db      599  ENRAFQRMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWS 658
Qy      659  RVMWDSKATLGLGGVAVNLGAVMAAGPFSYLGIRSSILVIOVDFVLVSGADNIFPV 718
Db      659  RVMWDSKATLGLGGVAVNLGAVMAAGPFSYLGIRSSILVIOVDFVLVSGADNIFPV 718
Qy      719  LEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTMPAVRTFALTSG 778
Db      719  LEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTMPAVRTFALTSG 778
Qy      779  AVILDFLLQMSAFVALLSLSKQBSARLDVCCCVKQDELPPPGQEGGLLGGFFOKAYAP 838
Db      779  AVILDFLLQMSAFVALLSLSKQBSARLDVCCCVKQDELPPPGQEGGLLGGFFOKAYAP 838
Qy      839  FLHWHITRGVV-----LLFLFALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLNRYF 894
Db      839  FLHWHITRGVVPSQLLFLFALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLNRYF 898
Qy      895  EVGAPVYFVTLGYNFSSEAGMNAICSSAGCNPFSTOKIOYATEFPFQSYLAIPASSWV 954
Db      899  EVGAPVYFVTLGYNFSSEAGMNAICSSAGCNPFSTOKIOYATEFPFQSYLSLPASSWV 958
Qy      955  DDFDMLTPSSCCRLYISGPNKDFCPTVNSLNCNKCMSITMGSVRPSVEQPHKLPW 1014
Db      959  DDFDMLTPSSCCRLYISGPNKDFCPTVNSLNCNKCMSITMGSVRPSVEQPHKLPW 1018
Qy      1015  FLNDRPNIKCPKGGAAVSTSVNLSDGQVLASRFMAHYPKPNQSDYTEALRAARELAA 1074
Db      1019  FLNDRPNIKCPKGGAAVSTSVNLSDGQVLASRFMAHYPKPNQSDYTEALRAARELAA 1078
Qy      1075  NITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTPEAVSCLLGLD 1134
Db      1079  NITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTPEAVSCLLGLD 1138
Qy      1135  LRSGLLNLSIYMLVDTVGFMAIDISYNAVSLINLSAVGMSVEFVSHITRSPAISTK 1194
Db      1139  LRSGLLNLSIYMLVDTVGFMAIGISYNAVSLINLSAVGMSVEFVSHITRSPAISTK 1198
Qy      1195  PTLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRMLNLTILGLLHG 1254
Db      1199  PTLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFPFRMLNLTILGLLHG 1258
Qy      1255  LVFLPVILSYGVDNVPALALBQKRAEEA-----VAAMVVASCPNHPSPRVST 1301
Db      1259  LVFLPVILSYGVDNVPALALBQKRAEEGGSGSHGGLAQIITPPSPQLTTSKST 1313

```

RESULT 34

```

PCT-US03-22467-2
; Sequence 2, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
PCT-US03-22467-2

```

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
 Best Local Similarity 77.8%; Pred. No. 0;
 Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

```

Qy      1  MABAGRLWLLALLRLAQSEPYTHIOPGYCAFYDECCKNPBLSGLMTLSNVCSLSN 60
Db      1  MAAAWL-GWLLWALLLSAAQGLTTPRHEAGVCTFYDECCKNPBLSGLTSLSNVCSLSN 59
Qy      61  TPARKITGDHLILLOKI CPRLYTGPNTQ- ACCSAKQLVSLBASISITKALITRCPACSDN 119
Db      60  TPASHVTGEHLALLQRI CPRLYNGPNTTFACCSTKQLLSLESSMSITKALLTRCPACSDN 119
Qy      120  FVNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEOSYDSCSRVRVPA 179
Db      120  FVSLHCHTCSPDQSLFINVTRVVERGAGEPPAVVAEAFYQHSFAEAYSCSVRIIPA 179
Qy      180  AATLAVGTMCGVYSALCNAQRLNFOGDTNGNGLAPLDTITPHLLPQAVGSGIGPLNEG 239
Db      180  AASLAVGSMCGVYSALCNAQRLNFOGDTNGNGLAPLDTITPHLLPQAVGSGIGLOPLNGK 239
Qy      240  VARNCSQGDVATCSCQDCAASCPAIARPOALDSTFYLQMPGSLVLIILCSFVAVVT 299
Db      240  IAPCNESQGDSDSVCSCQDCAASCVPVIPPPEALRPSFYMGRMPGWLIIITPAVFVLLS 299
Qy      300  ILLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFQFGQGTWVASHPLTILVLS 359
Db      300  AVLVELRVVSNRNKKAEGPQEAFLPHKHKLSPHTILGRFFQNMGTTRVASHPLTILVLS 359
Qy      360  VIPVVALAAGLVFTELTTPVELWSAPNSQARSEKAPHDOHGFPPFRTNQVILTAPNRSS 419
Db      360  FIVVVALAAGUTFIELTTTPVELWSAPNSQARSEKAFHDEHGFPPFRTNQVILTAPNRSS 419
Qy      420  YRYDSLLGPNKFGSGLDLDLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPN 479
Db      420  YKYDSLLGSKNFSGLSLDFLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPN 479
Qy      480  TSLDYCCINSLLQYFQNNRTLLLTANQTLNGQTSQVDKQHFLYCANAPLTFKQGTALA 539
Db      480  TSLDCCCVNLSLQYFQNNRTLLLTANQTLNGQTSQVDKQHFLYCANAPLTFKQGTALA 539
Qy      540  LSCMADYCAPVFPFLAIGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFL 599
Db      540  LSCMADYCAPVFPFLAVGQYQTDYSEAEALITFSLNNYPAGDPRLAQAKLWEAEFL 599
Qy      600  MRAFORMMAGMFOVTFTAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWSR 659
Db      600  MESFORNTSKFQVAFSAERSLEDEINRTTQDLVPFAVSYIIFLYISLALGSYSRCSR 659
Qy      660  VMWDSKATLGLGGVAVNLGAVMAAGPFSYLGIRSSILVIOVDFVLVSGADNIFPV 719
Db      660  VAVESKATLGLGGVAVNLGAVMAAGPFSYLGIRSSILVIOVDFVLVSGADNIFPV 719
Qy      720  EYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEALCFFLGALTMPAVRTFALTSG 779
Db      720  EYQRLPRRPGPREVHIGRTLGSVAPSMMLCSLSEALCFFLGALTMPAVRTFALTSG 779
Qy      780  VILDFLLQMSAFVALLSLSKQBSARLDVCCCVKQDELPPPGQEGGLLGGFFOKAYAP 839
Db      780  IILDFLLQMTAFVALLSLSKQBSARPDVLCCTFTRKLPPEKEGGLLRRFRKIYAP 839
Qy      840  LLHWITRGVWLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLNRYFVGP 899
Db      840  LLHREIRPVWMLLFTLFANLYLNCINVGDLQDELALPKDSYLLDYFLNRYLEVGP 899
Qy      900  VYFVTTLGYNFSSEAGMNAICSSAGCNPFSTOKIOYATEFPFQSYLAIPASSWVDDFD 959
Db      900  VYFVTTSFNFSSSEAGMNAICSSAGCKSFSLTKIOYASEFPDQSYVAIASSWVDDFD 959
Qy      960  WLTP-SSCCRLYISGPNKDFCPTVNSLNCNKCMSITMGSVRPSVEQPHKLPWFLND 1018
Db      960  WLTPSSCCRLYIRGPHKDFCPTDTSFNCNKCNWRTLGPVPTAEQPHKLPWFLND 1019
Qy      1019  RNNIKCPKGGAAVSTSVNLSDGQVLASRFMAHYPKPNQSDYTEALRAARELANITA 1078
Db      1020  PNRIKCPKGGAAVSTSVNLSDGQVIAQSFMAHYPKPNQSDYTEALRAARELANITA 1079
Qy      1079  DLKRVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTPEAVSCLLGLD 1138

```


; APPLICANT: Garcia-Calvo, Margarita
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF
; FILE REFERENCE: 36134-PCT 074669.0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; PRIOR FILING DATE: 2005-01-25
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
PCT-US05-01469-2

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFVDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGLYTPKHEAGVCTFYEECGKNPELSGSLMTLSNVSCLSN 59

QY 61 TPARKITGDHLILLOKICPRLYTGNTQ-ACCSAKQLVSLASLITKALLTRCPACSDN 119
DB 60 TPARKITGDHLILLOKICPRLYTGNTQ-ACCSAKQLVSLASLITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSFLNTRVAQIGAGQLPAVVAAYEAFYQHSFABQSDYSCSRVPA 179
DB 120 FVNLCHNTCSNQSFLNTRVAQIGAGQLPAVVAAYEAFYQHSFABQSDYSCSRVPA 179

QY 180 AATLAVGTGCVYGGALCNAQRLNFGQDTGNGLAPLDTFHLHPGQAVGSGIOPNEG 239
DB 180 AATLAVGTGCVYGGALCNAQRLNFGQDTGNGLAPLDTFHLHPGQAVGSGIOPNEG 239

QY 240 VARNESQDDVATCSQDCAASCPAIPALPQALDSTFYLGQMPGSLVLIILCSFVAVT 299
DB 240 IAPCNESQDDSAVCSQDCAASCPVIPPALPSPFYMGPMGLALIIITFAVFVLLS 299

QY 300 ILLVGFVAPADKSNWDPKGTSLSDKLSFTHTLLGQFQCGWGTWASHPLIIVLS 359
DB 300 AVLVRVVRNKNKAEGFQAPKLPHKHLSPHTILGRFFQNMGTFRVASWPLTVLALS 359

QY 360 VTPVVALAAGLFTLTDPVLSAPNSQARSEKAFDOHGFPPERTNOVILTAPNRSS 419
DB 360 FIVVIALAAGLFTLTDPVLSAPNSQARSEKAFDOHGFPPERTNOVILTAPNRSS 419

QY 420 YRYSLLGPKNFSGLDLDLLELELQERLRLHQLVNSPEAQRNLSQDICYAPLNPDN 479
DB 420 KYDLSLLGSKNFSGLDLDLLELELQERLRLHQLVNSPEAQRNLSQDICYAPLNPN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVDWKDHFLYCANAPLTFKDGTLA 539
DB 480 TSLSDCCVNSLLQYFQNNRTLLLTANQTLMGTSQVDWKDHFLYCANAPLTFKDGTLA 539

QY 540 LSCMADYGAPVPFPLAIGYKGVSEARALITWFSLNYPAGDPLRAQAKLWEAFLE 599
DB 540 LSCMADYGAPVPFPLAIGYKGVSEARALITWFSLNYPAGDPLRAQAKLWEAFLE 599

QY 600 MRAFORRMAGMQVTFABRSLEDEINRTTAEPLPFIATSYIVFLYISLALGSYSNR 659
DB 600 MESFORNTSKQVAFSAERSLEDEINRTTAEPLPFIATSYIVFLYISLALGSYSNR 659

QY 660 VMVDSKATILGGLGAVVILGAVMAAGFPFYLGRSSILVLOVFPFLVLSVGADNFI 719
DB 660 VAVESKATILGGLGAVVILGAVMAAGFPFYLGRSSILVLOVFPFLVLSVGADNFI 719

QY 720 EYQRLPRRCPREPHIGHALGRVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 779
DB 720 EYQRLPRRCPREPHIGHALGRVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 779

QY 780 VILDFLLQMSAFVALLSLDSKRQASRLDVCCVKPQELPPPGQSGLLGPFQKAYAPP 839
DB 780 VILDFLLQMSAFVALLSLDSKRQASRLDVCCVKPQELPPPGQSGLLGPFQKAYAPP 839

DB 780 IILDFLLQMSAFVALLSLDSKRQASRLDVCCVKPQELPPPGQSGLLGPFQKAYAPP 839
QY 840 LLHWITRGVLLLLFLALPGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGP 899
DB 840 LLHRFIRVVMVLLFLTLFGANLYMCNINVGLDQELALPKDSYLLDYFLFLNRYFEVGP 899
QY 900 VYFVTTGLGYNFSSBAGMNAICSSAGCNFSTFKIQIYATEPPEQSYLAIPASSWVDD 959
DB 900 VYFVTTGLGYNFSSBAGMNAICSSAGCNFSTFKIQIYATEPPEQSYLAIPASSWVDD 959
QY 960 WLTP-SSCCRLYISGPNKDFCPSTVNSLNCILKNCMSITMGSVRSVRSQFHKYLPWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDFCPSTVNSLNCILKNCMSITMGSVRSVRSQFHKYLPWFLND 1019
QY 1019 RPNIKCPKGGKLAAYSTSVNLTSDGOVLASRMAVHKPLKNSQDYTEALRAARELANITA 1078
DB 1020 PPNRCPKGGKLAAYSTSVNLTSDGOVLASRMAVHKPLKNSQDYTEALRAARELANITA 1079
QY 1079 DLKVPKGTDPAPFVPPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
DB 1080 DLKVPKGTDPAPFVPPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRS 1139
QY 1139 LNLISIVMILVDTVGFMAWDISVNAVSLINLVSAGMSVEFVSHITRSPAIKPTWL 1198
DB 1140 LNLISIVMILVDTVGFMAWDISVNAVSLINLVSAGMSVEFVSHITRSPAIKPTWL 1199
QY 1199 ERAKEATISMSGSAVAGVAMTNLPGILVLGLAKAQLIOIFFRNLNLTLLGLLHGLVFL 1258
DB 1200 ERAKDATVFMGSAVAGVAMTNLPGILVLGLAKAQLIOIFFRNLNLTLLGLLHGLVFL 1259
QY 1259 PVILSYVGPDPVNPALALRQKRAEAAVAVMVASCPHPSRVSTADNIVYNHSEFSI-KG 1317
DB 1260 PVILSYVGPDPVNPALALRQKRAEAAVAVMVASCPHPSRVSTADNIVYNHSEFSI-KG 1316
QY 1318 AGATSNFLPNNGRQP 1332
DB 1317 ANAARSLPKSDQKF 1331

RESULT 37

US-10-621-758A-2
; Sequence 2, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2

Query Match 78.5%; Score 5421.5; DB 36; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFVDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGLYTPKHEAGVCTFYEECGKNPELSGSLMTLSNVSCLSN 59

QY 61 TPARKITGDHLILLOKICPRLYTGNTQ-ACCSAKQLVSLASLITKALLTRCPACSDN 119
DB 60 TPARKITGDHLILLOKICPRLYTGNTQ-ACCSAKQLVSLASLITKALLTRCPACSDN 119

QY 120 FVNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
DB 120 FVSLHCHNTCSPDQSLFINVTRVERGAGEPPAVVAYEAFYQHSFAEKAYESCSQVRIIPA 179
QY 180 AATLAVGTWCGYVGSALCNAQWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEG 239
DB 180 AASLAVGSMCGYVGSALCNAQWLNFGQDTGNGLAPLDITFHLLEPGQALPDGIQPLNGK 239
QY 240 VARNESQGDVATCSQDCQCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSVPVAVT 299
DB 240 IAPCNESQGDSDAVCSQDCQCAASCPVIPPPEALRPSFYNGRMPGWLALIIITFAVFLVS 299
QY 300 ILLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPCQGWGTWVASWPLTILVLS 359
DB 300 AVLRLRVVNRNKNKAEGPOEAPKLPKHKLSPHTILGRPFQNGMTRVWASWPLTILVLS 359
QY 360 VTPVVALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHFGPPFRNTQVILTPANRNS 419
DB 360 FIVVIALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHFGPPFRNTQVILTPANRNS 419
QY 420 YRYDSLGLGPKNFSGILDLDLLELLELQERLRLHQLQVWSPAEARNISLQDICYAPLNPN 479
DB 420 YKYDSLGLGPKNFSGILDLDLLELLELQERLRLHQLQVWSPAEARNISLQDICYAPLNPN 479
QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVQWQKDFLYCANAPLTFKQGTALA 539
DB 480 TSLSDCCVNSLLQYFQNNRTLLLTANQTLMGQTSQVQWQKDFLYCANAPLTFKQGTALA 539
QY 540 LSCMADYGAPVPFELAIQYKGVKGYSEAEALTWTSLLNNYPAGDPRLAQAKLWEAFLE 599
DB 540 LSCMADYGAPVPFELAIQYKGVKGYSEAEALTWTSLLNNYPAGDPRLAQAKLWEAFLE 599
QY 600 MEAFQRMAGMFQVTFABRSLEDEINRTTARDLPFATSYIVIFLYISLALGSYSSWSR 659
DB 600 MESFORNTSDKQVAFSAERSLEDEINRTTIQDLPVFAVSIIYFVLYISLALGSYSSWSR 659
QY 660 VMWDSKATLGLGVAVVLGAVMAAGFFSYLGRSLVILQVVPFVLVSUGADNIFIVL 719
DB 660 VAVESKATLGLGVAVVLGAVMAAGFFSYLGRSLVILQVVPFVLVSUGADNIFIVL 719
QY 720 EYQRLPRRGEPREVHIGRALGRVAPSMLLCSLSAEICFPLGALTPMPAVRTFALTSGLA 779
DB 720 EYQRLPRRGEPREVHIGRALGRVAPSMLLCSLSAEICFPLGALTPMPAVRTFALTSGLA 779
QY 780 VILDFLLQMSAFVALLSLDSKQESRLDCCVCPQELPPPGQEGGLLGFQKAYAPF 839
DB 780 IILDFFLQMTAFVALLSLDSKQESRPDLCCPSTRKLPFPKXEGGLLRRFRKIYAPF 839
QY 840 LLHWITRGVLLFLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGP 899
DB 840 LLHRTFIRPVVMLFLTLFGANLYLMCNINVLQDQELALPKDSYLLDYFLFLNRYFEVGP 899
QY 900 VYFVTILGYNFSSBAGMNAICSSAGCNFSTQKIQYATEPPEQSYLAI PASSWDDFD 959
DB 900 VYFVTILGYNFSSBAGMNAICSSAGCNFSTQKIQYATEPPEQSYLAI PASSWDDFD 959
QY 960 WLTP--SSCRLVYISGNPKDPCPSTVNSLNCNLCNMSITWGSVRPSVQGFHYLWPLFND 1018
DB 960 WLTPSSCCRLVYIRGHKDEFCPSTDTSCNLCNLCNMTLGRVPRPTEQGFHYLWPLFND 1019
QY 1019 RPNIRCPKGLAAYSTSVNLSDGQVLAASFMAHYHKLKNSODYTEALRAARELANITA 1078
DB 1020 RPNIRCPKGLAAYSTSVNLSDGQVLAASFMAHYHKLKNSODYTEALRAARELANITA 1079
QY 1079 DLKVPQGTDPAPVEVPYITNIFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSG 1138
DB 1080 DLKVPQGTDPAPVEVPYITNIFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSG 1139
QY 1139 LNLLSIVMLVDYTGFWALMDIISYNAVSLINLVSAGVMSVFEVSHITRSPAIKPTWL 1198
DB 1140 LNLLSIIMILVDYTGFWALMDIISYNAVSLINLVSAGVMSVFEVSHITRSPAIKPTWL 1199
QY 1199 ERAKEATISMGSAVFAGVAMTNPGLVLGLAKAQLIQIFFRLNLLITLLGLLHGLVFL 1258

DB 1200 ERAKQATVPMGSANFAGVAMTNPFGILILGPAQAQLIQIFFRLNLLITLLGLLHGLVFL 1259
QY 1259 FVILSVGVDNPNPALALEOKRAEEAAVAAVMVASCPNHPBSRVSTADNIYVNHSPSGSI-KG 1317
DB 1260 FVILSVGVDNPNPALALEOKRAEEAAVAAVMVASCPNHPBSRVSTADNIYVNHSPSGSI-KG 1316
QY 1318 AGAINSFILPNNGRQF 1332
DB 1317 ANAARSSILPKSDQF 1331
RESULT 38
US-10-646-301A-2
; Sequence 2, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-646-301A-2
Query Match 78.5%; Score 5421.5; DB 36; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
QY 1 MRAEGLRGWLLWALLRLAQSEPYTIIHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQSGELYTPKHEAGVCTFYECGKNPELSGGITLSNVSCLSN 59
QY 61 TPARKITGDHILLOKICPRLVYTGNTQ-ACCSAQVLSLEASLSITTKALLTRCPACSN 119
DB 60 TPARKITGDHILLOKICPRLVYTGNTQ-ACCSAQVLSLEASLSITTKALLTRCPACSN 119
QY 120 FVNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
DB 120 FVSLHCHNTCSPDQSLFINVTRVERGAGEPPAVVAYEAFYQHSFAEKAYESCSQVRIIPA 179
QY 180 AATLAVGTWCGYVGSALCNAQWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEG 239
DB 180 AASLAVGSMCGYVGSALCNAQWLNFGQDTGNGLAPLDITFHLLEPGQALPDGIQPLNGK 239
QY 240 VARNESQGDVATCSQDCQCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSVPVAVT 299
DB 240 IAPCNESQGDSDAVCSQDCQCAASCPVIPPPEALRPSFYNGRMPGWLALIIITFAVFLVS 299
QY 300 ILLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPCQGWGTWVASWPLTILVLS 359
DB 300 AVLRLRVVNRNKNKAEGPOEAPKLPKHKLSPHTILGRPFQNGMTRVWASWPLTILVLS 359
QY 360 VTPVVALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHFGPPFRNTQVILTPANRNS 419
DB 360 FIVVIALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHFGPPFRNTQVILTPANRNS 419
QY 420 YRYDSLGLGPKNFSGILDLDLLELLELQERLRLHQLQVWSPAEARNISLQDICYAPLNPN 479
DB 420 YKYDSLGLGPKNFSGILDLDLLELLELQERLRLHQLQVWSPAEARNISLQDICYAPLNPN 479
QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVQWQKDFLYCANAPLTFKQGTALA 539

```
Db 480 TSLSDCCVNSLLQYFQNNRTLLMTANQTNGQTSVDMKDHFLYCANAPLTFKDGTSLA 539
Qy 540 LSCMADYGAPVFPFLAIGYKGYKGYSEABALIMTFSLNNYPAGDPRLAQAOKLWEBAFLBE 599
Db 540 LSCMADYGAPVFPFLAIGYKGYKGYSEABALIMTFSLNNYPAGDPRLAQAOKLWEBAFLBE 599
Qy 600 MRAFORMMAGFQVTFPFAERSLEDEINRTTAEEDLPIFATSYIVIFLYISLALGSYSWSR 659
Db 600 MESFORNTSDKPFQVAFSAERSLEDEINRTTQDLFVFAVSIIYFVLYISLALGSYSRCSR 659
Qy 660 VMVDSKATLGLGGVAVVLGAVMAAGFFSILGIRSSVLQVVPFLVLSVGDNIFIFVL 719
Db 660 VAVESKATLGLGGVAVVLGAVLAAAGFYSLGVPSSLVIQVVPFLVAVGADNIFIFVL 719
Qy 720 EYQRLPRPGEPREVHIIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Db 720 EYQRLPRPGEPREVHIIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Qy 780 VILDPELLQMSAFVALLSLDSKQREASRLDVCCVCKPQELPPGQEGILLGFQKAYAPF 839
Db 780 VILDPELLQMTAFVALLSLDSKQREASRPDLVCCFSTRKLPDPPEKEGGLLFRFKIYAPF 839
Qy 840 LLMHTIRGVLILFLALRGVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRYEFGAP 899
Db 840 LLHRIRPVMLLFTLFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGP 899
Qy 900 VYFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFID 959
Db 900 VYFVTTSFNFSEAGMNAICSSAGCKSFSLTKIYASEFPDQSYVAIAASSWDDFID 959
Qy 960 WLTP-SSCCRLYISGPNKDKFCPSTVNSLNCILKNCMSITMGSVRSVQFHKYLPWFLND 1018
Db 960 WLTPSSCCRLYIRGPHKDFECPSTDTSPNCILKNCNRTLGEVRSPTAQFHKYLPWFLND 1019
Qy 1019 RNIKCPKGLAAYSTSVNLSDGQVLASRFMAYHKPLKNSODYTEALRAARELANITA 1078
Db 1020 PNIRKCPKGLAAYSTSVNLSDGQVIAQFVAYHKPLNSQDTEALRASLLAANITA 1079
Qy 1079 DLKRVGTDPAFEVFPYITINVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
Db 1080 DLKRVGTDNPEVFPYITINVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDMCSG 1139
Qy 1139 LNLASIVMLVDYTFQFMAWDIISYNAVSLINLVSAGVSVFVSHITRSPASTKPTWL 1198
Db 1140 ILNLSIIMLVDTITGLMAVWGISYNAVSLINLVTAVGMSVFEVSHITRSPAVTKPTKL 1199
Qy 1199 ERAKEATISMGSAVFAGVAMTNLPGILVLGAKAQLIQIFFFFRLMLLITLLGLLHGLVFL 1258
Db 1200 ERAKATVPMGSVAVGAVAMTNFPGILILGFAQAQLIQIFFFFRLMLLITLLGLLHGLVFL 1259
Qy 1259 PVILSYGPDVNPALALEOKRAEAAVAAVMVASCNPHPSRVSTADNIYVNHSPFEGSI-KG 1317
Db 1260 PVVLSYLGPDVNOALVOEKLASEA-AVAPEPSCQYPSPADAN--VNYGFAPELAHG 1316
Qy 1318 AGAISNPLNNGROP 1332
Db 1317 ANAARSSLPKSDQKF 1331
```

RESULT 39

```
US-10-663-208A-2
; Sequence 2, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
```

```
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-663-208A-2
```

```
Query Match 78.5%; Score 5421.5; DB 36; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAEAGRLGWLWALLLRLAQSEPYTHIQPYCAYFYDCGKNPELGSIMTLNVSCLSN 60
Db 1 MAEAGRLGWLWALLLRLAQSEPYTHIQPYCAYFYDCGKNPELGSIMTLNVSCLSN 59
Qy 61 TPARKITGDHLLILKICPRLYTGTNTQ-ACCSAKQLVLSLEASLITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILKICPRLYTGTNTQ-ACCSAKQLVLSLEASLITKALLTRCPACSDN 119
Qy 120 FVNLHCHNTCSFNOSLFINVTRVAQAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
Db 120 FVNLHCHNTCSFNOSLFINVTRVAQAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
Qy 180 AATLAVGTCMGVYSALCNAQRWLNFGQDTNGNLAPLDITFHLLEPQAVSGGIQPLNEG 239
Db 180 AATLAVGTCMGVYSALCNAQRWLNFGQDTNGNLAPLDITFHLLEPQAVSGGIQPLNEG 239
Qy 240 VARNESGDDVATCSCQDCAASCPAIAPQALDSTFVLGOMPSLVLIIILCSYPAVVT 299
Db 240 VARNESGDDVATCSCQDCAASCPAIAPQALDSTFVLGOMPSLVLIIILCSYPAVVT 299
Qy 300 ILLVGFVAPARDKSKMVDPKGTSLSDKLSFTHTLLGQFFQGMGTWVASMPLTILVLS 359
Db 300 ILLVGFVAPARDKSKMVDPKGTSLSDKLSFTHTLLGQFFQGMGTWVASMPLTILVLS 359
Qy 360 VIVVALAAGLVTFELTATDPVELMSAPNSQARSEKAFHQHFGPFPFRNQVLTAPNRSS 419
Db 360 VIVVALAAGLVTFELTATDPVELMSAPNSQARSEKAFHQHFGPFPFRNQVLTAPNRSS 419
Qy 420 YRYSLLILGPKNFSGILDLLELLELEQERLHLQVMSPEAQRNLSQDICYAPLNPDN 479
Db 420 YRYSLLILGPKNFSGILDLLELLELEQERLHLQVMSPEAQRNLSQDICYAPLNPDN 479
Qy 480 TSLYDCCINSLLQYFQNNRTLLMTANQTNGQTSVDMKDHFLYCANAPLTFKDGTSLA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLMTANQTNGQTSVDMKDHFLYCANAPLTFKDGTSLA 539
Qy 540 LSCMADYGAPVFPFLAIGYKGYKGYSEABALIMTFSLNNYPAGDPRLAQAOKLWEBAFLBE 599
Db 540 LSCMADYGAPVFPFLAIGYKGYKGYSEABALIMTFSLNNYPAGDPRLAQAOKLWEBAFLBE 599
Qy 600 MRAFORMMAGFQVTFPFAERSLEDEINRTTAEEDLPIFATSYIVIFLYISLALGSYSWSR 659
Db 600 MRAFORMMAGFQVTFPFAERSLEDEINRTTAEEDLPIFATSYIVIFLYISLALGSYSWSR 659
Qy 660 VMVDSKATLGLGGVAVVLGAVMAAGFFSILGIRSSVLQVVPFLVLSVGDNIFIFVL 719
Db 660 VAVESKATLGLGGVAVVLGAVLAAAGFYSLGVPSSLVIQVVPFLVAVGADNIFIFVL 719
Qy 720 EYQRLPRPGEPREVHIIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Db 720 EYQRLPRPGEPREVHIIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Qy 780 VILDPELLQMSAFVALLSLDSKQREASRLDVCCVCKPQELPPGQEGILLGFQKAYAPF 839
Db 780 VILDPELLQMTAFVALLSLDSKQREASRPDLVCCFSTRKLPDPPEKEGGLLFRFKIYAPF 839
```

QY 840 LLHWITRGVLLILFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGP 899
DB 840 LLHFRIRPVVMLLFLTLFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGP 899
QY 900 VYFVITLVNFSSEAGNNAICSSAGCNPFSTOKIOYATFPEQSYLAIPASSWVDDFID 959
DB 900 VYFVITLVNFSSEAGNNAICSSAGCKSFSLTOKIQVASEFPQSQVVAIAASSWVDDFID 959
QY 960 WLTP-SSCCRLYISGPNKOKPCPSTVNSLNCNCHSITMGSVRPSVEQPHKYLWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCLKNCNRTLGPVPTAEQPHKYLWFLND 1019
QY 1019 RNKICPKGGLAAYSTSVNLTSQVLAASHFMAHYHPLKNSQDYTEALRAARELANITA 1078
DB 1020 PPNIRCPKGLAAYSTSVNLSSDQVIAQFMAHYHPLKNSQDYTEALRAARELANITA 1079
QY 1079 DLKRVGTDPAPFVFPVYITNVFYEQVLTTLPEGLPMLSCLVPTFAVSCLLGLDLRS 1138
DB 1080 DLKRVGTDPFVFPVYITNVFYEQVLTTLPEGLPMLSCLVPTFAVSCLLGLDLRS 1139
QY 1139 LNLISVIMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198
DB 1140 LNLISVIMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1199
QY 1199 ERAKEATISGSAFVAGVAMTNLPGLVLGLAKAQILQIFFFRLNLLITLGLLHGLVFL 1258
DB 1200 ERAKDATVFMGSAFVAGVAMTNLPGLVLGLAKAQILQIFFFRLNLLITLGLLHGLVFL 1259
QY 1259 PVLTSVGVDPNPALALEOKRAEAAVAVVASCNPHSRVSTADNINYNHSPESI-KG 1317
DB 1260 PVLTSVGVDPNPALALEOKRAEAAVAVVASCNPHSRVSTADNINYNHSPESI-KG 1317
QY 1318 AGAISNLFNPNRGQF 1332
DB 1317 ANAARSSLPKSDQKF 1331

RESULT 40

US-10-736-769-2

Sequence 2, Application US/10736769

GENERAL INFORMATION:

APPLICANT: Altmann, Scott W

APPLICANT: Wang, Luquan

APPLICANT: Graziano, Michael

APPLICANT: Murgolo, Nick

TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF

FILE REFERENCE: J01603-K3-US

CURRENT APPLICATION NUMBER: US/10/736,769

CURRENT FILING DATE: 2003-12-16

PRIOR APPLICATION NUMBER: 60/397,442

PRIOR FILING DATE: 2002-07-19

PRIOR APPLICATION NUMBER: 10/621,758

PRIOR FILING DATE: 2003-07-17

PRIOR APPLICATION NUMBER: 10/646,301

PRIOR FILING DATE: 2003-08-22

PRIOR APPLICATION NUMBER: 10/663,208

PRIOR FILING DATE: 2003-09-16

NUMBER OF SEQ ID NOS: 51

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 1331

TYPE: PRT

ORGANISM: Rattus sp.

US-10-736-769-2

Query Match 78.5%; Score 5421.5; DB 37; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLIRLAQSEPTTIHQPGYCAFDECKRNPGLSGSLMTLSNVSCLSN 60
DB 1 MAAAML-GWLLWALLSAAQGLYTPKRBAGVCTFYBECGRNPGLSGSLMTLSNVSCLSN 59

QY 61 TPARKITGDHLLILLQKICPRLYTGPNQ-ACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
DB 60 TPARKITGDHLLILLQKICPRLYTGPNQ-ACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
QY 120 FVNLHCHNTCSNQSIFINVTVAQIAGQOLPAVAYEAPYQHSFABEQSVDSCSRVRVPA 179
DB 120 FVNLHCHNTCSNQSIFINVTVAQIAGQOLPAVAYEAPYQHSFABEQSVDSCSRVRVPA 179
QY 180 AATLVGTMGVGVYSALCNAQWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEG 239
DB 180 AATLVGTMGVGVYSALCNAQWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEG 239
QY 240 VARNESQGDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIIILCSVFAVVT 299
DB 240 IAPCNESQGDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIIILCSVFAVVT 299
QY 300 ILLVGRVAPADPKSNVDPKGTSLSDKLSSTHTLLGQFFQCGWGTWASWPLTILVLS 359
DB 300 AVLRLRVVSNRNKNAEGQAPKLPKHKLSPHTILGRFFQNGWTRVASWPLTILVLS 359
QY 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGPFPFRTQVILTAPNRSS 419
DB 360 FIVVIALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGPFPFRTQVILTAPNRSS 419
QY 420 YRYSLLLGKPNFSGILDLDLLELELERLHLQVWSPQARNISLQDICYAPLNPDN 479
DB 420 YKYDSLLLGKPNFSGILDLDLLELELERLHLQVWSPQARNISLQDICYAPLNPDN 479
QY 480 TSLYDCCINSILQYFQNNRTLLLTANQTLMGTSQVNDKDHFLYCANAPLTFKDGTALA 539
DB 480 TSLDCCVNSLLQYFQNNRTLLLTANQTLMGTSQVNDKDHFLYCANAPLTFKDGTALA 539
QY 540 LSCMADYGAPVFPFLAIGYKGDYSEARALTWPSLNYPAGDPLAQAQKWEAPLKE 599
DB 540 LSCMADYGAPVFPFLAIGYKGDYSEARALTWPSLNYPAGDPLAQAQKWEAPLKE 599
QY 600 MRAPQRMAGMFQVTFPERSLEDEINRTTAEDELPIFATSYIVIFLYISALGSYSWSR 659
DB 600 MESFORNTSDKQVAFSAERSLEDEINRTTQDLPVFAVSYIVIFLYISALGSYSWSR 659
QY 660 VMVDSKATLGLGVAVVILGAVMAAGFYSYLGIRSLVILQVVPFVLVSVGADNIIFVL 719
DB 660 VAVESKATLGLGVAVVILGAVMAAGFYSYLGIRSLVILQVVPFVLVSVGADNIIFVL 719
QY 720 EYQRLPRRGPBPREVHIGRALGRVAPSMILCSLSRAICPFLGALTMPAVRTFALTSGLA 779
DB 720 EYQRLPRRGPBPREVHIGRALGRVAPSMILCSLSRAICPFLGALTMPAVRTFALTSGLA 779
QY 780 VILDELLQMSAFVALLSLDSKQEAERLDVCCVCPQELPPQGGEGLLGFPQKAYAPF 839
DB 780 IILDFFLQMTAFVALLSLDSKQEAERLDVCCVCPQELPPQGGEGLLGFPQKAYAPF 839
QY 840 LLHWITRGVLLILFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGP 899
DB 840 LLHFRIRPVVMLLFLTLFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGP 899
QY 900 VYFVITLVNFSSEAGNNAICSSAGCNPFSTOKIOYATFPEQSYLAIPASSWVDDFID 959
DB 900 VYFVITLVNFSSEAGNNAICSSAGCKSFSLTOKIQVASEFPQSQVVAIAASSWVDDFID 959
QY 960 WLTP-SSCCRLYISGPNKOKPCPSTVNSLNCNCHSITMGSVRPSVEQPHKYLWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCLKNCNRTLGPVPTAEQPHKYLWFLND 1019
QY 1019 RNKICPKGGLAAYSTSVNLTSQVLAASHFMAHYHPLKNSQDYTEALRAARELANITA 1078
DB 1020 PPNIRCPKGLAAYSTSVNLSSDQVIAQFMAHYHPLKNSQDYTEALRAARELANITA 1079
QY 1079 DLKRVGTDPAPFVFPVYITNVFYEQVLTTLPEGLPMLSCLVPTFAVSCLLGLDLRS 1138
DB 1080 DLKRVGTDPAPFVFPVYITNVFYEQVLTTLPEGLPMLSCLVPTFAVSCLLGLDLRS 1139
QY 1139 LNLISVIMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198


```
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US03-22467-12

Query Match      78.3%; Score 5407; DB 1; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGLLWALLRLAQSEPTTIHQPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA 63
DB 3 AAQOGLLWALLNSAQGLYPTTHKAGCTFYECGKNPELSGLMTLSNVSCLSNTPA 62

QY 64 RKITGDHLLLOKICRPLYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSNFVN 122
DB 63 RHVTGDHLLLOKICRPLYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSNFVS 122

QY 123 LHCHNTCSNQSLFINVTRVAQAGOLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAT 182
DB 123 IHCHNTCSNQSLFINVTRVQDQGLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAS 182

QY 183 LAVGTGCVGYGALCNAQRLWLPQDGTGNGLAPLDTIFHLHPQAVGSGIOPLENGVAR 242
DB 183 LAVGSMCGVYGALCNAQRLWLPQDGTGNGLAPLDTIFHLHPQAVGSGIOPLENGVAR 242

QY 243 CNEQGDVVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILCSFVAVVTILL 302
DB 243 CNEQGEDSAACSQDCAACSPVIPPALRFSFTNGRMPGMLALIIITFAVFLSVVL 302

QY 303 VGRVAPARDKSMVDPKKGTSLSDKLSFTHTLLGQFGWGTWVASWPLTILVLVSIP 362
DB 303 VYLRVASNKNKTAGSQAPNLPRKRFSPTVLGRFPESNGTRVASWPLTILVLSFV 362

QY 363 VVALAAGLVPTELTDVPLWAGAPNSQARSEKAFHQDHPGPPFRNQVILTPAKNSRYK 422
DB 363 VIALSVGLTIFELTIDVPLWAGAPNSQARSEKAFHQDHPGPPFRNQVILTPAKNSRYK 422

QY 423 DSIILGPKNFGSLDLDLLELLEQLERHLQVMSPEAQRNLSLODICYAPLNPDNTSL 482
DB 423 DSIILGPKNFGSLDLDLLELLEQLERHLQVMSPEAQRNLSLODICYAPLNPDNTSL 482

QY 483 YDCINSLLOQYQNNRTLLLTANOTLMQTSQVMDKHFLYCANAPLTFKDGTALALSC 542
DB 483 TDCCVNSLLQYQNNHTLLLTANQTLNGQTSQVMDKHFLYCANAPLTFKDGTALALSC 542

QY 543 MADYGAPVPPFLAIGYKGYKDYSEBALIMTFSNLNYPAGDPRLOAKLWEEAFLEEMRA 602
DB 543 IADYGAPVPPFLAIGYKGYKDYSEBALIMTFSNLNYPAGDPRLOAKLWEEAFLEEMRA 602

QY 603 FORRMAGMEQVTFATERSLEDRNTTADLPIFATSYIVIFLYISLALGSSYSSSRVWV 662
DB 603 FORSTADKEQIAPSAERSLEDRNTTQDLPVFAISLYIVIFLYISLALGSSYSSSRVAV 662

QY 663 DSKATILGLGVAVLGVAVMAAGFFSYLGIRSLVILQVVPFLVLSVGDADNIFIFVLEYQ 722
DB 663 DSKATILGLGVAVLGVAVMAAGFFSYLGIRSLVILQVVPFLVLSVGDADNIFIFVLEYQ 722

QY 723 RLPRPGEPREVHIHGRALGRVAPSMMLCSLSBAICFFLGLALTPMPAVRTPALTSLGLVIL 782
DB 723 RLPRPGEPREVHIHGRALGRVAPSMMLCSLSBAICFFLGLALTPMPAVRTPALTSLGLVIL 782

QY 783 DFLQMSAFVALLSIDSKEQASRLDVCVQPELPPPGQEGILLGFFQKAYAPFLH 842
DB 783 DFLQMTAFVALLSIDSKEQASRLDVCVQPELPPPGQEGILLGFFQKAYAPFLH 842

QY 843 WITRGVLLFLALFGVSLWSCHISVGLDQELALPKDSYLLDYFLNRYPEVGPVYF 902
DB 843 RIRPVLLFLVLFANLYLMCNISVGLDQDIALPKDSYLLDYFLNRYLEVGPPVYF 902

QY 903 VTTLGYNFSSEAGMNAICSSAGCNFNFSTQKIYATEFPPEQSYLAIPASSWVDIDWLT 962
```

RESULT 43

```
PCT-US03-40113-12
; Sequence 12, Application PC/TUS0340113
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-WI
; CURRENT APPLICATION NUMBER: PCT/US03/40113
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US03-40113-12
```

```
Query Match      78.3%; Score 5407; DB 1; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGLLWALLRLAQSEPTTIHQPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA 63
DB 3 AAQOGLLWALLNSAQGLYPTTHKAGCTFYECGKNPELSGLMTLSNVSCLSNTPA 62

QY 64 RKITGDHLLLOKICRPLYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSNFVN 122
DB 63 RHVTGDHLLLOKICRPLYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSNFVS 122

QY 123 LHCHNTCSNQSLFINVTRVAQAGOLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAT 182
DB 123 IHCHNTCSNQSLFINVTRVQDQGLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAS 182
```


Db 603 FORSTADKFOIARSABRSLEDEINRTTIQDLPVFAISYLIVFYISLALGSGSRWGRVAV 662
QY 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGRSSVLIVQVVPFLVSVGADNIFIFVLEYQ 722
Db 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGRSSVLIVQVVPFLVSVGADNIFIFVLEYQ 722
QY 723 RLPRRPGREHIGRALGRVAPSMMLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMFGREAHIGRTLGSVAPSMMLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
QY 783 DFLQMSAFVALLSLDSKQESRLDVCCVCKPQELPPCQCEGLLGFQKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQESRLDVCCVCKPQELPPCQCEGLLGFQKAYAPFLH 842
QY 843 WITRGVLLLFALFGVSYLSMCHISVGLDQELALPKDSYLLDYFLFLNRYEFGVAPVYF 902
Db 843 RPIRPVLLLFVLFGANLYLNCNISVGLDQELALPKDSYLLDYFLFLNRYEFGVAPVYF 902
QY 903 VTTLYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT 962
Db 903 DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIQYASEFPNQSYVAIAASSWVDDFIDWLT 962
QY 963 P-SSCCRLYISGPNKDKFCPTVNSLNCNCKMSITWGSVRSPVEQFHKYLPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDEFCEPSTDTSPNCLKNMRTLGPRVPTTEQFHKYLPWFLNDRN 1022
QY 1022 IKCPKGLAAYSTVNLTSQVLAASRMAHYHKLKNSQDYTEALRAARELANITADLR 1081
Db 1023 IKCPKGLAAYSTVNLSDGQIIASQFMAHYHKLKNSQDYTEALRAARELANITADLR 1082
QY 1082 KVPGETDPAFVPPYTTINVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLN 1141
Db 1083 KVEGTDNPEVPPYTTISNVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLN 1142
QY 1142 LLSIVMLVDTGFMALNDISYNVSLINLVSAVMSVEFVSHITRSPAIKPTWLER 1201
Db 1143 LLSIIMLVDTGFMALNDISYNVSLINLVSAVMSVEFVSHITRSPAIKPTWLER 1202
QY 1202 KEATISMSAGVAVANTNLPGLVLGLAKAOLIQIFFPRLNLLITLGLLHGLVPLVI 1261
Db 1203 KDATIFMSAGVAVANTNFPGLILGFAQAOLIQIFFPRLNLLITLGLLHGLVPLVI 1262
QY 1262 LSYVGDVNPALALEOKRAEAAVAVMVASCPNHPRSRSTADNIYVNHSPFEGS-1KGAGA 1320
Db 1263 LSYVGDVNPALALEOKRAEAAVAVMVASCPNHPRSRSTADNIYVNHSPFEGS-1KGAGA 1321
QY 1321 ISNPLPNRQF 1332
Db 1322 ASSSLPKSQK 1333

RESULT 45
PCT-US05-27579-3
; Sequence 3, Application PC/TUS0527579
; GENERAL INFORMATION:
; APPLICANT: Mount Sinai School of Medicine
; APPLICANT: Ioannou, Yiannis
; APPLICANT: Davies, Joanna P.
; TITLE OF INVENTION: NPC1L1 AND NPC1L1 INHIBITORS AND METHODS OF USE THEREOF
; FILE REFERENCE: 2201581-WO0
; CURRENT APPLICATION NUMBER: PCT/US05/27579
; CURRENT FILING DATE: 2005-08-12
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US05-27579-3

Query Match

78.3%; Score 5407; DB 1; Length 1333;

Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGWLLWALLLELAQSEPYTTIHQPGYCAFYDECGKPELSGLMTLSNVSCLSNTPA 63
Db 3 AAQGGWLLWALLLNSAQGELYTTHKAGCTFYBECGKPELSGGTSLNVSCLSNTPA 62
QY 64 RKITGDHLLLOKICPRLYTGN-TOACCSAQVLSLEASLSITKALLTRCPACSNFVN 122
Db 63 RHVTGDHLLALQVCPRLYNGPNDYVACCSAQVLSLEASLSITKALLTRCPACSNFVN 122
QY 123 LHCHNTCSNQSLFINVTRVAQAGQLPAVYAEFYQHSFAEQSYDSCSRVRVPAAT 182
Db 123 IHCHNTCSNQSLFINVTRVQORDPQQLPAVYAEFYQHSFAEQSYDSCSRVRVPAAT 182
QY 183 LAVGTMCVYGSALCNAORWLNFGQDTGNGLAPLDTFHLLRPGQAVGSGIQPLNREGVAR 242
Db 183 LAVGTMCVYGSALCNAORWLNFGQDTGNGLAPLDTFHLLRPGQAVGSGIQPLNREGVAR 242
QY 243 CNESQDDVATCSQDCAASCPAIAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTIL 302
Db 243 CNESQEDSAAACSCQDCAASCPVIPPALRPSFYNGRMPGWLALIIITAVFVLSVVL 302
QY 303 VGRVAPARDKSMVDPPKGTSLSDKLSFTHTLLGQFPQGMGTWVASPLTILVLSVLP 362
Db 303 VYLRVASNRKNKTAGSQEAPNLPRKRPSPTVLGRPFESWGTRVASWPLTVLALSFI 362
QY 363 VVALAAGLVFTLTTPDVELMSAPNSOARSEKAFHDOHGFPPRTNOVILTAPNRSRY 422
Db 363 VIALSVGLTFTLTTPDVELMSAPNSOARSEKAFHDOHGFPPRTNOVILTAPNRSRY 422
QY 423 DSLLLGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 482
Db 423 DSLLLGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLEL 482
QY 483 YDCCINSLLQYFONNRTLLLTANQTLMGQTSQVMDKHFLYCANAPLTFKDGTALALSC 542
Db 483 TDCCVNSLLQYFONNHTLLLTANQTLMGQTSQVMDKHFLYCANAPLTFKDGTALALSC 542
QY 543 MARYGAPVPPFLAIGYKGYKDYSEBALIWTSLNYPAGDPRLAQAKIWEERAFLEEMRA 602
Db 543 IADYGAPEPFLAVGQYQGTDYSEBALIITFSINYPADDPMAHAKIWEERAFLEEMQS 602
QY 603 FORRMAGMFQVTFTAERSLEDEINRTTADLPFATSYIVIFLYISLALGSGSRWGRVAV 662
Db 603 FORSTADKFOIARSABRSLEDEINRTTIQDLPVFAISYLIVFYISLALGSGSRWGRVAV 662
QY 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGRSSVLIVQVVPFLVSVGADNIFIFVLEYQ 722
Db 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGRSSVLIVQVVPFLVSVGADNIFIFVLEYQ 722
QY 723 RLPRRPGREHIGRALGRVAPSMMLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMFGREAHIGRTLGSVAPSMMLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
QY 783 DFLQMSAFVALLSLDSKQESRLDVCCVCKPQELPPCQCEGLLGFQKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQESRLDVCCVCKPQELPPCQCEGLLGFQKAYAPFLH 842
QY 843 WITRGVLLLFALFGVSYLSMCHISVGLDQELALPKDSYLLDYFLFLNRYEFGVAPVYF 902
Db 843 RPIRPVLLLFVLFGANLYLNCNISVGLDQELALPKDSYLLDYFLFLNRYEFGVAPVYF 902
QY 903 VTTLYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT 962
Db 903 DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIQYASEFPNQSYVAIAASSWVDDFIDWLT 962
QY 963 P-SSCCRLYISGPNKDKFCPTVNSLNCNCKMSITWGSVRSPVEQFHKYLPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDEFCEPSTDTSPNCLKNMRTLGPRVPTTEQFHKYLPWFLNDRN 1022
1022 IKCPKGLAAYSTVNLTSQVLAASRMAHYHKLKNSQDYTEALRAARELANITADLR 1081

Db 1023 IRCPKGGLAAVYRTSVNLSSDQGIASQFMAVYHKPLRNSQDTEALRASLLAANTAE LR 1082
Qy 1082 KVPGTDPAPFVFPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLIN 1141
Db 1083 KVPGTDPNPFVFPYTTISNVFYQQYLTVLPEGFTLALCFVPTFVVCYLLGLDLRSGILN 1142
Qy 1142 LLSIWMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSFAISKPTWLE RA 1201
Db 1143 LLSIWMILVDTITGLMAVWGISYNAVSLINLVTAVGMSVEFVSHITRSFAVSTKPTRLERA 1202
Qy 1202 KEATISMSGSAFAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLTLLGLLHGLVFLPVI 1261
Db 1203 KDATIFMSGSAFAGVAMTNFPGILLGLGFAQAQLIQIFFFRNLNLTLLGLLHGLVFLPV 1262
Qy 1262 LSYGPDVNPALALQKRAEAAVAVMVASCPNHPFSRVSTADNIYVNHSPFGS- IKGAGA 1320
Db 1263 LSYLPGDVNQALVLEKLATEA- AMVSEBSCPYFPFADANTSDYVNYGFNPEFPEINA 1321
Qy 1321 ISNPLPNNGRQF 1332
Db 1322 ASSSLPKSDQKF 1333

RESULT 46

US-10-621-758A-12

; Sequence 12, Application US/10621758A

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan

; APPLICANT: Graziano, Michael

; APPLICANT: Murgolo, Nick

; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF

; FILE REFERENCE: JBO1603-K-US

; CURRENT APPLICATION NUMBER: US/10/621,758A

; CURRENT FILING DATE: 2003-07-17

; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 1333

; TYPE: PR

; ORGANISM: Mus sp.

US-10-621-758A-12

Query Match

Best Local Similarity 78.3%; Score 5407; DB 36; Length 1333;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

Qy 4 AGRGWLWALLRLAQSPEYTTIHQPGYCAFYDCGKNPELSGSIWTLNVSCLSNTPA 63
Db 3 AAQGWLLWALLNSAQSGLYTPTHKAGCTFYEBGKNPELSGGLTSLNISCUSNTPA 62
Qy 64 RKITGDHLTLLOKICPLRYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSDNFVN 122
Db 63 RHVTGDHLLALORVCPRLYNGNDTYACCSTKQLVSLDSSLSITKALLTRCPACSENFS 122
Qy 123 LHCHNTCSNQSLFINVTRVAGLQALPAVAYEAFYQHSFABQSDYSCSRVRVPAAT 182
Db 123 IHCHNTCSQSLFINVTRVQRPQALPAVAYEAFYQHSFABQSDYSCSRVRVPAAS 182
Qy 183 LAVTGCMGYGSAALCNAQRLWLFQDGTGNGLAPLDTFHLBPQGVGSGIOPANEV 242
Db 183 LAVTGCMGYGSAALCNAQRLWLFQDGTGNGLAPLDTFHLBPQGVGSGIOPANEV 242
Qy 243 CNESQDDVATCSQDCAASCPAIAARPOALDSTFYLGQMPGSLVLIILCSVFVAVTILL 302
Db 243 CNESQEDSAAACSCQDCAASCPVIPPALRSFYNGRMFGWLAIIIFTAVVLLSVL 302
Qy 303 VGRVAPADKSKMDPKKGTSLSDKLSFTHTLQGFQGGWGTWVASWPLTILVLSVP 362
Db 303 VTLRVASNRNKNKTAGSQAPNLPKRKRPSPHTVLGRFFESWGTRVASWPLTILVLSFV 362

RESULT 47

US-10-646-301A-12

; Sequence 12, Application US/10646301A

; GENERAL INFORMATION:

Qy 363 VVALAAGLVFTLTTPVELMSAPNSQARSEKAPHDQHFPGFFRTNOVILTAPNRSSRY 422
Db 363 VIALSVGLTFTLTTPVELMSAPNSQARSEKAPHDQHFPGFFRTNOVILTAPNRSSRY 422
Qy 423 DSSLGPKNFSGILLDLLELLELQERLRHLQVMSPEAQNIISLQDI CYAPLNDPNTSL 482
Db 423 DSSLGPKNFSGILLDLLELLELQERLRHLQVMSPEAQNIISLQDI CYAPLNDPNTSL 482
Qy 483 YDCCINSLQYFQNNRTLLITANOTLMGOTSQVDMKDHFLYCANAPLTFKDGTAALASC 542
Db 483 TDCCVNSLQYFQNNHTLLITANOTLNGOTSLVDMKDHFLYCANAPLTYKDGTAALASC 542
Qy 543 MADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLWBEAFLEEMRA 602
Db 543 IADYGAPVFPFLAVGGYQGTDYSEAEALITFSNNYPADDPRAHAKLWBEAFLEKMQS 602
Qy 603 FORMMAGMFQVTFABRSLEDEINRTTABDLPPIFATSYIVIFLYISLALGYSYSSRWV 662
Db 603 FORSTADKQFAFSAERSLEDEINRTTQDLPVFAISYLIIVFLYISLALGYSYSSRWV 662
Qy 663 DSKATILGGLGVAVVLGAVMAAGFYSYLGVPSSLVIIQVVPFLVLAVGADNIFIFVLEYQ 722
Db 663 DSKATILGGLGVAVVLGAVMAAGFYSYLGVPSSLVIIQVVPFLVLAVGADNIFIFVLEYQ 722
Qy 723 RLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFGLGALTMPMPAVRTFALTSGIAVIL 782
Db 723 RLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFGLGALTMPMPAVRTFALTSGIAVIL 782
Qy 783 DFLQMSAFVALSLDSKRQBSRLDVCCKVKPQBELPPPGQEGLLGLFPQKAYAPFLILH 842
Db 783 DFLQMTAFVALLSLDSKRQBSRPDVVCCFSRRNLPPPKQEGLLGLCFRRKIYFPFLILH 842
Qy 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLPLNRYFEVGPVVF 902
Db 843 RFIIPVLLFLVLFGANLYLWCNISVGLDQDLPKDSYLLDYFLPLNRYFEVGPVVF 902
Qy 903 VTTLYNFSBAGMNAICSSAGCNFSTQKIQTAFEPFQSYLAIPASSWDDFIDWLT 962
Db 903 DTTSGYNFSTAGMNAICSSAGCSFSLTQKIYASEFPNQSYVAIAASSWDDFIDWLT 962
Qy 963 P-SSCCRLYISGPNKDKFCPTVNSLNCIKCMSTMGSVRPSVQFHKYLPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDEFPCPTDTSFNCLKNCMRTLGPVTRPTTQFHKYLPWFLNDRN 1022
Qy 1022 IKCPKGLAAVYSTVNLTSQVLAASRFMAVYHKPLKNSQDYTEALRAARELAANTADLR 1081
Db 1023 IRCPKGGLAAVYRTSVNLSSDQGIASQFMAVYHKPLRNSQDTEALRASLLAANTAE LR 1082
Qy 1082 KVPGTDPAPFVFPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLIN 1141
Db 1083 KVPGTDPNPFVFPYTTISNVFYQQYLTVLPEGFTLALCFVPTFVVCYLLGLDLRSGILN 1142
Qy 1142 LLSIWMILVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSFAISKPTWLE RA 1201
Db 1143 LLSIWMILVDTITGLMAVWGISYNAVSLINLVTAVGMSVEFVSHITRSFAVSTKPTRLERA 1202
Qy 1202 KEATISMSGSAFAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLTLLGLLHGLVFLPVI 1261
Db 1203 KDATIFMSGSAFAGVAMTNFPGILLGLGFAQAQLIQIFFFRNLNLTLLGLLHGLVFLPV 1262
Qy 1262 LSYGPDVNPALALQKRAEAAVAVMVASCPNHPFSRVSTADNIYVNHSPFGS- IKGAGA 1320
Db 1263 LSYLPGDVNQALVLEKLATEA- AMVSEBSCPYFPFADANTSDYVNYGFNPEFPEINA 1321
Qy 1321 ISNPLPNNGRQF 1332
Db 1322 ASSSLPKSDQKF 1333

```

; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-646-301A-12

Query Match      78.3%; Score 5407; DB 36; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGHLLALLRLAQSEPTTHIQPGYCAFYDECGKPELSGLMTLSNVSLCNTPA 63
DB 3 AAQGGHLLALLNSAQSELYTTHKAGFCFTFEECGKPELSGLMTLSNVSLCNTPA 62
QY 64 RKITGPHLLILLOKICPLRYTGNP-TOACCSAKQLVLEASLSITKALLTRCPACSNFVN 122
DB 63 RHVTGPHALLQVCEPLRYNGPNDTYACCSKQLVLEASLSITKALLTRCPACSNFVS 122
QY 123 LHCHNTCSNQSIFINVTVAQAGQLPAVAYEAFYQHSFAEQSYDCSRVVPAAAT 182
DB 123 IHCHNTCSNQSIFINVTVAQAGQLPAVAYEAFYQHSFAEQSYDCSRVVPAAAT 182
QY 183 LAVGTGCGYGSALCNAQRLWLFQGTGNGLAPLDTIFHLLPQAGVSGIOPNLEGVAR 242
DB 183 LAVGSMCGYGSALCNAQRLWLFQGTGNGLAPLDTIFHLLPQAGVSGIOPNLEGVAR 242
QY 243 CNESQDDVATCSQDCAACPAIARPOALDSTFYLQMGPGSILVLIILCSFVAVTILL 302
DB 243 CNESQDDVATCSQDCAACPAIARPOALDSTFYLQMGPGSILVLIILCSFVAVTILL 302
QY 303 VGRVAPARDKGMVDPKGTSLSDKLSFTSTHLLQFFQGGTGWASWPLTILVLSVP 362
DB 303 VYLRVASNENKNTAGSQEAPNLPRKRFSPHTVLGRFFESWGTRVASWPLTILVLSFV 362
QY 363 VVALAAGLVFTEITDTPVELWAPNSQARSEKAFHQHGFPPFRNQVILTAPNRSYRY 422
DB 363 VIALSVGLTFIELTDPVELWAPNSQARSEKAFHQHGFPPFRNQVILTAPNRSYRY 422
QY 423 DSILLGPKNFGSLDLDLLELELELERHLQVMSPEAQRNLSLODICVAPLNDTSL 482
DB 423 DSILLGPKNFGSLDLDLLELELELERHLQVMSPEAQRNLSLODICVAPLNDTSL 482
QY 483 YDCINSLLOQYQNNKRLTLLTANOTLMQTSQVQWKDHFLYCANAPLTFKDGTALALSC 542
DB 483 TDCCVNSLLQYQNNHTLLTANQTLNGQTSILVDKQDHFLYCANAPLTFKDGTALALSC 542
QY 543 MADYGAPVFPFLAIGYKGYSEABALIMTFSLLNYPAGDPRLAQAKLWEEAFLEEMRA 602
DB 543 IADYGAPVFPFLAIGYKGYQGTDYSEABALITFSINNYPADDPMAHAHLWEEAFLEEMQS 602
QY 603 FORMAGMPOVTFEATERSLEDEINRTAEDLIPATSYIVIFLYTSLALGSSYSSRWVW 662
DB 603 FORMAGMPOVTFEATERSLEDEINRTAEDLIPATSYIVIFLYTSLALGSSYSSRWVW 662
QY 663 DSKATLGLGGVAVLVGAVMAAGFFSYLGRSSVLQVVPFLVSVGADNIPFVLEQY 722
DB 663 DSKATLGLGGVAVLVGAVMAAGFFSYLGRSSVLQVVPFLVSVGADNIPFVLEQY 722
QY 723 RLPRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGGALTPMPAVRTFALTSGLAIVL 782

```

RESULT 48

```

US-10-663-208A-12
; Sequence 12, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-663-208A-12

```

```

Query Match      78.3%; Score 5407; DB 36; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

```

```
QY 4 AGLRGWLLWALLRLRLAQSEPYTHIOPGYCAFYDECGKNPELSSGLMTLSNVSCLSNTPA 63
DB 3 AAQOGWLLWALLLNSAQGELYTPTHKAGFCTFYEECGKNPELSSGLTSLNSISCLSNTPA 62
QY 64 RKITGDHLLILOKICPRLYTGN-TQACCSAKQLVLEASLSITKALLTRCPACSDNFVN 122
DB 63 RHVTGDHLLALQVCPRLYNGPNDIYACCSKQLVSLDSSLITKALLTRCPACSENFS 122
QY 123 LHCHNTCSNQLSFLNVTNRVAQLGAGOLPAVVAYEAFYQHSFABOSYDSCRVRVPAAT 182
DB 123 IHCHNTCSPDQSLFNVTRVQDQGLPAVVAYEAFYQHSFABOSYDSCRVRVPAAT 182
QY 183 LAVGTWCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPQOAVGSGIOPNEGVAR 242
DB 183 LAVGSMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPQOAVGSGIOPNEGVAR 242
QY 243 CNESQDDVATCSQDCAASCPAIAARPQALDSTFYLGQWPGSLVLIILCSVFAVVTILL 302
DB 243 CNESQEDSAAACSCQDCAASCPVIPPALRPSFYMGPMGLALIIIFTAVFVLLSVUL 302
QY 303 VGRVAPARDKMDPKKGTSSDKLSKLSFSTHTLLIGQFFQGGTWWASWPLTILVLSVIP 362
DB 303 VYLRVASNRNKNKTAGSQEAPNLPKRKRPSPTHVLRPFESGTRVASWPLTVLALSIV 362
QY 363 VVALAAGLVFTETLTPDVELWAPNSQARSSEKAFHQHGGPPERTNOVILLTAPNESSYRY 422
DB 363 VVALSVGLTFIEITDTPVELWAPNSQARSSEKAFHQHGGPPERTNOVILLTAPNESSYRY 422
QY 423 DSSLGPKNFSIGILDLILLELELEQERLHIVQMSPEAQRNISLQDICYAPLNDTSL 482
DB 423 DSSLGPKNFSIGILDLILLELELEQERLHIVQMSPEAQRNISLQDICYAPLNDTSL 482
QY 483 YDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVNDKOHFLYCANAPLTFKDGTLALASC 542
DB 483 TDCCVNSLLQYFQNNHTLLLTANOTLNGQTSQVNDKOHFLYCANAPLTFKDGTLALASC 542
QY 543 MARYGAPVPELAIQYKGYKQYSEALITWTSNNYPAGDPRLAQAKLWEAFLEEMBA 602
DB 543 IADYGAPVFPFVAVGQYQDTSSEALITWTSNNYPAGDPRLAQAKLWEAFLEEMBA 602
QY 603 FORMMAGMPQVFTFAERSLEDEINRTTARDLPIFATSYIVIFLYISLALGSYSSSRVAV 662
DB 603 FORSTADKEQIAFSAERSLEDEINRTTQDLVFAISLYIVFLYISLALGSYSSSRVAV 662
QY 663 DSKATLGLGGVAVVLGAVMAANGFYSLGVPSLVIIQVVPFLVLAAGADNIFIFVLEYQ 722
DB 663 DSKATLGLGGVAVVLGAVMAANGFYSLGVPSLVIIQVVPFLVLAAGADNIFIFVLEYQ 722
QY 723 RLPRRGPBREVHIQALGRVAPSMILLCSLSAICFFELGALTPMAVRTFALTSLGLAVIL 782
DB 723 RLPRMPGEQREAHIGTLGSAVPSMILLCSLSAICFFELGALTPMAVRTFALTSLGLAVIL 782
QY 783 DFLQMSAFVALLSLDSKQEASRLDVCCCVKQPELPPPGQEGELILGPFQKAYAPFLIH 842
DB 783 DFLQMTAFVALLSLDSKQEASRPDVCCVPSRNLPPPKQEGELILGPFQKAYAPFLIH 842
QY 843 WTRGVVLLFLALFVGSILYSWCHISVGLDQELAPKDSYLLDYFLFLNRYEFGAPVYF 902
DB 843 RFRPVPVLLFLVLFQANLYLNCNISVGLDQELAPKDSYLLDYFLFLNRYEFGAPVYF 902
QY 903 VTTLVNPSSEAGMNAICSSAGCNFESFTKIQYATEPPEQSVLAI PASSSWDDFDWILT 962
DB 903 DTTSGYVNFSTEAGMNAICSSAGCESFSLTKIQYASEFPNQSVLA PASSSWDDFDWILT 962
QY 963 P-SSCCRLYISGPNKDFCPSTVNSLNCNKMISITWGSVRPSVEQFHYLPFWLNDREN 1021
DB 963 PSSCCRIYTRGPHKDFECPSTDTSFNCLNKNMRTLGPVPTTEQFHYLPFWLNDTEN 1022
QY 1022 IKCPKGLLAAYSTSVNLTSDGQVLAQFMAHYHKLKNSQDYTEALRAARELANITADUR 1081
DB 1023 IRCPKGLLAAYRTSVNLSDDQIIASQFMAHYHKLKNSQDYTEALRAARELANITAEUR 1082
```

```
QY 1082 KVPGTDPAFEPFPTITNVFYEQLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLLN 1141
DB 1083 KVPGTDPNFEPFPTISNVFYQQYLVTLPEGIFTLALCFVPTFVVCYLLGLDIRSGILN 1142
QY 1142 LLSIVMILVDTVGPMAWLDISYNAVSLINLVASVGMSEFVSHITRSPAIKPTWLER 1201
DB 1143 LLSIIMILVDTIGLWVGISYNAVSLINLVAVGMSVEFVSHITRSPAVSTKPTRLBRA 1202
QY 1202 KEATISGSAVPAVGMNTNLPGLVLGLAKAQLQIIFFRNLNLTLLGLLHGLVFLPVI 1261
DB 1203 KDAITFMSGSAVPAVGMNTNFPGLILGLPAQQLQIIFFRNLNLTLLGLLHGLVFLPV 1262
QY 1262 LSYVGPDPNPALEQKRAEBAVAAMVASCNPHPSRVSTADNIIYVNSHFECS-1KGAGA 1320
DB 1263 LSYLGPDPNPALEQKRAEBAVAAMVASCNPHPSRVSTADNIIYVNSHFECS-1KGAGA 1321
QY 1321 ISNFLPNNGROF 1332
DB 1322 ASSSLPKSDQKF 1333
```

RESULT 49

```
US-10-736-769-12
; Sequence 12, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K3-US
; CURRENT APPLICATION NUMBER: US/10736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRN
; ORGANISM: Mus sp.
US-10-736-769-12
```

```
Query Match 78.3%; Score 5407; DB 37; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;
```

```
QY 4 AGLRGWLLWALLRLRLAQSEPYTHIOPGYCAFYDECGKNPELSSGLMTLSNVSCLSNTPA 63
DB 3 AAQOGWLLWALLLNSAQGELYTPTHKAGFCTFYEECGKNPELSSGLTSLNSISCLSNTPA 62
QY 64 RKITGDHLLILOKICPRLYTGN-TQACCSAKQLVLEASLSITKALLTRCPACSDNFVN 122
DB 63 RHVTGDHLLALQVCPRLYNGPNDIYACCSKQLVSLDSSLITKALLTRCPACSENFS 122
QY 123 LHCHNTCSNQLSFLNVTNRVAQLGAGOLPAVVAYEAFYQHSFABOSYDSCRVRVPAAT 182
DB 123 IHCHNTCSPDQSLFNVTRVQDQGLPAVVAYEAFYQHSFABOSYDSCRVRVPAAT 182
QY 183 LAVGTWCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPQOAVGSGIOPNEGVAR 242
DB 183 LAVGSMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPQOAVGSGIOPNEGVAR 242
QY 243 CNESQDDVATCSQDCAASCPAIAARPQALDSTFYLGQWPGSLVLIILCSVFAVVTILL 302
DB 243 CNESQEDSAAACSCQDCAASCPVIPPALRPSFYMGPMGLALIIIFTAVFVLLSVUL 302
```


[illegible]

```

RESULT 51
PCT-US01-08631-53050
; Sequence 53050, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53050
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-53050

```

Query Match	64.6%;	Score 4466;	DB 1;	Length 982;
Best Local Similarity	98.0%;	Pred. No. 0;		
Matches 867;	Conservative 1;	Mismatches 9;	Indels 8;	Gaps 3;
QY	1	MAEAGLRGWLWALLRLAQSPPYTIHOGVCYAFYDECGKNPELGGSLMTLNSVCLSN	60	
DB	1	MAEAGLRGWLWALLRLAQSPPYTIHOGVCYAFYDECGKNPELGGSLMTLNSVCLSN	60	
QY	61	TPARKITGDHLILLQKICPRLYTGPNTQACCSAKQLVSLSEASISITKALLTRCPACSDNF	120	
DB	61	TPARKITGDHLILLQKICPRLYTGPNTQACCSAKQLVSLSEASISITKALLTRCPACSDNF	120	
QY	121	VNLHCHNTCSPNQSFLINTRVAQLGAGQLPAVVAEAFYQHSFASQSDCSRVRPAA	180	
DB	121	VNLHCHNTCSPNQSFLINTRVAQLGAGQLPAVVAEAFYQHSFASQSDCSRVRPAA	180	
QY	181	ATLAVGTMCGVYGSALCNQARWLNFQDGTGNGLAPLDITFHILLEPCQAVGSGIQPLNEGV	240	
DB	181	ATLAVGTMCGVYGSALCNQARWLNFQDGTGNGLAPLDITFHILLEPCQAVGSGIQPLNEGV	240	
QY	241	ARCNESQGDVDATCSCQDCAASCPC--AIARPOALDSTFYLGQMPGSLVLIILICSVFVV	298	
DB	241	ARCNESQGDVDATCSCQDCAASCPCPAPRSPSTPW--RQWPGSLVLIILICSVFVV	298	
QY	299	TILLVGRVAPARDKSKVDPKGTSLSOKLSFSTHTLLGQPFQCHGTWVASWPLTILVL	358	
DB	299	TILLVGRVAPARDKSKVDPKGTSLSOKLSFSTHTLLGQPFQCHGTWVASWPLTILVL	358	
QY	359	SVTPVVALAAGLVFTELTTPVELWAPNSQARSEKAFHDQHGPPFRNTQVILTAPNRS	418	
DB	359	SVTPVVALAAGLVFTELTTPVELWAPNSQARSEKAFHDQHGPPFRNTQVILTAPNRS	418	
QY	419	SVRYDSLLGPKNFSGILDLLLELLELLEQLERHLQVMSPEAQRNISLQDICYAPLNPD	478	
DB	419	SVRYDSLLGPKNFSGILDLLLELLELLEQLERHLQVMSPEAQRNISLQDICYAPLNPD	478	
QY	479	NTSLYDCCINSLLQYQNNRTLLLTANOTLMGQTSQVDWKDHFLVCANAPLTFKGTAL	538	
DB	479	NTSLYDCCINSLLQYQNNRTLLLTANOTLMGQTSQVDWKDHFLVCANAPLTFKGTAL	538	
QY	539	ALSCMADYGAPVPPFLAIGYKGKDYSEAEALIMTFLSNYPAGDPRLAQAKLWEAEAPLE	598	
DB	539	ALSCMADYGAPVPPFLAIGYKGKDYSEAEALIMTFLSNYPAGDPRLAQAKLWEAEAPLE	598	
QY	599	EMEAFFORMAGMVFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSYSSWS	658	
DB	599	EMEAFFORMAGMVFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSYSSWS	658	
QY	659	RVMVDSKATILGIGGVAVVLGAVMAAGFFSYLGIRSLVLTVQVPPFLVLTVSGADNIIFFV	718	
DB	659	RVMVDSKATILGIGGVAVVLGAVMAAGFFSYLGIRSLVLTVQVPPFLVLTVSGADNIIFFV	718	
QY	719	LEYORLPRRPGEPREVHIGALGRVAPSMLLCSISRAICFFLGAITPMPAVRTPALTSG	778	
DB	719	LEYORLPRRPGEPREVHIGALGRVAPSMLLCSISRAICFFLGAITPMPAVRTPALTSG	778	
QY	779	AVTLDLFLQMSAPVALLSLDSKQEAESRLDVCCKVQPELPPPCQSGELLGFFQKAYAP	838	
DB	779	AVTLDLFLQMSAPVALLSLDSKQEAESRLDVCCKVQPELPPPCQSGELLGFFQKAYAP	838	
QY	839	FLLHWITRGVW----LLLFLALFGVLSYMSCHISVGLDQBELALPK	879	
DB	839	FLLHWITRGVWPSQLLLFLALFGVLSYMSCHISVGLDQBELALPK	883	

```

RESULT 52
US-10-450-763-53050
; Sequence 53050, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763

```



```
Db 473 PPLDQILHQLVDLQIAIEN--ITASYDNETVTLQDICIAPLSPYNT---NCTILSVLNY 527
Qy 494 FQNNRTLLLTANQTLMGQTSQVMDKHDFLYCANAPLTFKQGTALALSCMADYGAPVPPF 553
Db 528 FQNSHVLVDHKKGDDFF---VYADYTHFLYCVRAPASLNDTSLLDHPCFLGTGFGVFPFW 584
Qy 554 LAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEBEMRAFORRMAGMFQV 613
Db 585 LVLGYYDQYNNATLVITPPVNNYNDTEKLQRAQAWKEFNFVKNYKN---PNLTI 641
Qy 614 TFTAERSLEDEINRTAEDLPIFATSYIVIFLYISALGSYSSWRVMVDSKATLGLGV 673
Db 642 SFTAERSIEDELNRSDSVFTVWISYAIMFLYISALGHKSCRRLLDVKSLGIAGI 701
Qy 674 AVVLGAVMAAMGFFSYLGRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRPEPRE 733
Db 702 LIVLSSVACSLGVFSYIGLPLTVIEVIFLVAVGVNDNIFILVQAYQDERLQGETLD 761
Qy 734 VHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSLGLAVILDFLLQMSAFVA 793
Db 762 QQLGRVLGEVAPSMFLSFSFETVAFVLGALSVMPAVHTFSLFAGLAVFDLLOITCFVS 821
Qy 794 LLSLDSKQKEASRLDVCCVKPQELPPQ--GEGILLGFFQKAYAPFLLHMTIRGVLL 852
Db 822 LLGLDIKQKRNRLDIFCCVRAEDGTSVQASESCLFRFFKNSYSPLLLKQWNRPIVIAI 881
Qy 853 FLALFGVSLYSCHISVGLDQELALPKDSYLLDYFLNRYFEVGAUVYFVTLGYNFSS 912
Db 882 FVGVLFSIAVLNKVDIGLDQSLMPDDSYWYDFKSIQYLHAGPPYFVFLBEGHDYTS 941
Qy 913 EAGMAICSSAGCANNFSFTQIKQYATEPPEQSYLAIPASSWVDDFIDMLTP--SSCCRLYI 971
Db 942 SKGQNVCGMGCCNDSLVQOIFNAQAOLDNYTRIGFAPSSWIDDFYDWKQSSCCRV-- 999
Qy 972 SGNPKDKFCPSVNSLNCNCKMT--MGSVRPSVEQPHKYLPWFINDRPNIKCPKGGIA 1030
Db 1000 -DNITDQFCNASVDPACVR--CRPLTPGKQRPQGGDFMRFLPMLSDNPNPKCGKGHA 1057
Qy 1031 AVYSTN--LTSQGVLASREMAXHKLKNSODYTEALRAARELANITADLRKVPGETD 1088
Db 1058 AVSSAVNLLGHGTRVGYTFYTHVTLQTSADFIDALKARLIASNV--ETWNGINGS-- 1114
Qy 1089 AFEVFPYTTINVEYQYLTILPEGLFMLSCLVPTFAVSCLLLGLDLSGLNLLSIYMI 1148
Db 1115 AXYVFPYSYVYFYQYLTIIIDTFNLGVSIGAFLVYTWLLGCELWSAVIMCATIAMV 1174
Qy 1149 LVDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHITRSPATSKPTWLERAKATISM 1208
Db 1175 LVNMFQVMWLGISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRVERAEBALAHM 1234
Qy 1209 GSAVPAVAVMTNLPGLTILGLAKAQLIQIFFRLNLLITLGLLHGLVPLVLSVGPD 1268
Db 1235 GSSVFSGITLTKFGGIWLAFAKSQIFQIFYFRMYLAWYLLGATHGLPLPLVLLSYIGPS 1294
Qy 1269 VNPA 1272
Db 1295 VNKA 1298
```

RESULT 54

```
US-10-940-774A-10152
; Sequence 10152, Application US/10940774A
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774A
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10152
; LENGTH: 1318
; TYPE: PRN
; ORGANISM: Human
US-10-940-774A-10152
```

```
Query Match 34.8%; Score 2403; DB 39; Length 1318;
Best Local Similarity 39.7%; Pred. No. 2.2e-201;
Matches 525; Conservative 238; Mismatches 451; Indels 110; Gaps 26;
```

```
Qy 2 AEAGLRGWLWA-----LILRLAQSEPVYTIHQPGYCAFYDECG-----KNPELSG 47
Db 32 ASALRGHSMRTARGLAGLLGLLLCPAQVFSQ-----SCVMYGEGLIAYGDKRYNCEYSG 86
Qy 48 SLMTLSNVCSLSNTPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASLITK 107
Db 87 -----PPKPLPKDGYDLVQELCPGFFFG--NVSLCCDVRQLQTLKDNLQLPL 131
Qy 108 ALLTRCPACSNFVNLHCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAEYAFYQHS 163
Db 132 QFLSRPCSPFNLLNLFCELTCSPRSQFLNVTATEDYDVPVTNQTKTNVKELOYYVGS 191
Qy 164 FAEQSYDSCSRVRPAATLAVGTMCVYVGSALCNAQRLNLFQDGTGNGLAPLDIT--- 219
Db 192 FANAMYNACRDEVEAPSSNDKALGLLCKGADA--CNATNWEYFNKNDGQAPFTITPVS 250
Qy 220 ---PHLLPFGQAVSGIQIPLNEGVARCNESQGDVATCQDCAASCAPAIARPQ----- 270
Db 251 DFPVH-----GMEPMNATKGCDESDEVTAPCSCDCSIVCGPKQPPPPAPW 300
Qy 271 ---ALDSTFYLGQWPGSLVLIILCSFPAVNTILL-----VGRVAPARDKS 314
Db 301 TILGLDAMVYIMWITYMAFLVFFGAPFAVMCYRKRYFVSEYTPIDSNIAFSV--NASDKG 359
Qy 315 KMVDPKKTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLVSVIPVVALAAGLVFTE 374
Db 360 E-----ASCCDPVSAAFEGCLRLRFLTRWGSFCVRNPGCVIFFSLVITACSSGLVFR 412
Qy 375 LTTDPVELWAPNSQARSEKAFHDQFGPPFRFTNQVILITAPNRSYRYDSLLLGPK--NFS 433
Db 413 VTTNPVDLWASAPSSQARLEKEYFDQFGFFRTEQLIIRAPLTDKHIQYPSGADVPRG 472
Qy 434 GILDLDLLELLELQERLRHQLQWSPEAQRNISLQDICVAPLNPDNTSLDYDCINSLLQY 493
Db 473 PPLDQILHQLVDLQIAIEN--ITASYDNETVTLQDICIAPLSPYNT---NCTILSVLNY 527
Qy 494 FQNNRTLLLTANQTLMGQTSQVMDKHDFLYCANAPLTFKQGTALALSCMADYGAPVPPF 553
Db 528 FQNSHVLVDHKKGDDFF---VYADYTHFLYCVRAPASLNDTSLLDHPCFLGTGFGVFPFW 584
Qy 554 LAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEBEMRAFORRMAGMFQV 613
Db 585 LVLGYYDQYNNATLVITPPVNNYNDTEKLQRAQAWKEFNFVKNYKN---PNLTI 641
Qy 614 TFTAERSLEDEINRTAEDLPIFATSYIVIFLYISALGSYSSWRVMVDSKATLGLGV 673
Db 642 SFTAERSIEDELNRSDSVFTVWISYAIMFLYISALGHKSCRRLLDVKSLGIAGI 701
Qy 674 AVVLGAVMAAMGFFSYLGRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRPEPRE 733
Db 702 LIVLSSVACSLGVFSYIGLPLTVIEVIFLVAVGVNDNIFILVQAYQDERLQGETLD 761
Qy 734 VHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSLGLAVILDFLLQMSAFVA 793
Db 762 QQLGRVLGEVAPSMFLSFSFETVAFVLGALSVMPAVHTFSLFAGLAVFDLLOITCFVS 821
Qy 794 LLSLDSKQKEASRLDVCCVKPQELPPQ--GEGILLGFFQKAYAPFLLHMTIRGVLL 852
Db 822 LLGLDIKQKRNRLDIFCCVRAEDGTSVQASESCLFRFFKNSYSPLLLKQWNRPIVIAI 881
```



```
PCT-US03-40978-1542
; Sequence 1542, Application PC/TUS0340978
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: PCT/US03/40978
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1542
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40978-1542

Query Match      34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy      7  RGWLWALLRLRLAQSEPTTTHQPGYCAPYDECG-----KNPELSGLMTLSNVSCLS 59
Db      4  RGLALGLLILLCPAQVFSQ-----SCVMYGEIGAYGDKRYNCEYSG-----46

Qy      60  NTPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLRSLSTIKALLTRCPACSDN 119
Db      47  --PPKPLPKDGVLDVQELCPGFFFG-NVSLCCDVRQLQTLKDNQLQPLQFSRCSFCYN 103

Qy      120  FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYBAFYQHSFAEQSYDSCSRV 175
Db      104  LNLNLCFELTCSRQSOFLNVTATEDYDVPVNTQKTNVKELYYQGSFANAMYNACRDV 163

Qy      176  RVPAAATLAVGTMCYGYGALCNAQRWLNFGQDTGNGLAPLDIT-----PHLEPGQA 228
Db      164  BAPSSNDKALGLCGKDADA-CNATNWIEMFNKNGQAPFTITPVFSDFPVH-----215

Qy      229  VSGIQTPLNEGVARNESGDDVATCSQDCAASCPATAPQ-----ALDSTFYLG 279
Db      216  ---GHEPMNATKGDCEVDEVTAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272

Qy      280  OMPGSLVLIILCSVFAVVITLL-----VGRVPAPARDKSMVDPKKGTSL 326
Db      273  WITYWFLLVFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy      327  DKLSFSTHTLLOFGQGTWVASWPLTILVLSIPVVALAAGLVFTTELTDPVELWSAP 386
Db      325  DPVSAAFEGCLRLFRMGSCVRNPGCVIFFSLVFITACSSGLVFVRVTTNPVDLWSAP 384

Qy      387  NSQARSEKAFPHOHEGPPERTNOVILTAPNRSYRYSLLGPK-NFSGILDLDLLELL 445
Db      385  SSOARLEKEYFOHGFPPFRTTQELIRAPLTDKHIYQYPYSGADVPFGPPLDIQLHQL 444

Qy      446  ELQERLRHLQVMSPEAQRNISLQIDICYAPLNPNTSLYDCINSLLIQYFQNNRTLLTLTA 505
Db      445  DLQIAIEN--ITASYDNEVTIQLDCLAPLSYNT---NCTILSVLNYFQNSHVLDHKK 499

Qy      506  NQTLMGQTSQVDKWHFLYCANAPITFKDGTALALSCMADYGAIPVPPFLAIGYKGDYS 565
Db      500  GDDFF---VYADVHTFLYCVRAPASLNDTSLLDHDCILGTFGPGVPPMLVLGDDYDQYN 556

Qy      566  EBAALIMTSLNYPAGDPRLAQAKLWEAFIEEMRAPQRMAGMPQVTFPFAERSLEDEI 625
Db      557  NATALVITFPVNNYNDTEKLQRAQAWKEPINFVKNYKN---PNLTISFTAERSIEDSL 613

Qy      626  NRTTARDLPIFATSYIVIFLYTSLALGYSNRSVMVDKATLGLGGVAVLVGAVMAAG 685
Db      614  NRESDSVTVVISVAIMFLYISLALCHYKSCERLLVDKSVLSGIGAILIVLSSVACSLG 673

Qy      686  FFSYLGISSVLQVVPFLVLSVGADNIFIFVLEYQRLPRPGEPREVHI GRALGRVAP 745
Db      674  VFSYIGLPLTIVIEIPFLVLAAGVDNIFILVQAYORDERLQGETLDQQLGRVLGEVAP 733
```

```
Qy      746  SMLLCSLSEAICTFFLGALTMPAVRTFALTSGLAVILDFLLQMSAFVALLSLDSRQERAS 805
Db      734  SMFLSSSEFETVAFFLGALSVMVAHTTSLFAGLAVFIDFLQITCFVSLGLDIRQKRN 793

Qy      806  RLDVCCCKVQQLPPPGQ--GGLLLGPPQKAYAPLLHWITRGVVLVLLFLALFGVSLYGM 864
Db      794  RLDIFCCVRGAEDGTSTQASESCLFRFPKNSYSPLLKDKWMPVIAIFVGVLSFIAVL 853

Qy      865  CHISVGLDOELALPKDSYLLDYLFLNRYFVFGAPVYFVTTLYGNFSEAGNACSSAG 924
Db      854  NKVDIGLDQSLSDMPDDSYMDYFKSISQYLHAGPPVYFVLEBEGHDYTSKGGQNMVCCGMG 913

Qy      925  CNNEFSTQIKIYATEFPEQSYLAIPASSWVDDFDWLTG--SSCCRLYISGPNKDKFCPST 983
Db      914  CNNDLSLVQOIFNAQAQLDNYTRIGFAPSSWIDDDYDMVKPQSSCCRV---DNITDQFCNAS 970

Qy      984  VNSNLCKNCKMSIT-MGSVRPSVQFHKYLPWFNLDRNPKICPKGGLLAAYSTSVN--LTS 1040
Db      971  VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSSAVNILLGH 1029

Qy      1041  DGQVLASFPMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVPEYTTNV 1100
Db      1030  GTRVGATYFTMYHTVTLQTSADFIDALKKARLIASNVT-ETWINGS--AYRVFPGSVFY 1086

Qy      1101  FVEQVLTILPBGFLMFLSLCLVPTFAVSCLLGLDLRSGLLNLSTVIMILVDTVGFMAWD 1160
Db      1087  FVEQVLTITIDTIFNLGVSIGALFLVTVVLLGCELWSAVIMCATIAMVLVNMFGVWLWG 1146

Qy      1161  ISYNAVSLINLVSAGMSVFEVSHITRSPATSTKTWLERAKBATISMGSAVFAVAMTN 1220
Db      1147  ISLNAVSLVNLVMSGCGISVFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK 1206

Qy      1221  LPGILVLGLAKAQLIQIFPRLNLLITLGLLHGLVLPVLVILSYVGGPDVNP 1272
Db      1207  FGIIVVTLAPAKSQIPQIFYFMYLANVLLGATHGLIFLPLVLSYIGPSVNKA 1258

RESULT 57
PCT-US04-07141-810
; Sequence 810, Application PC/TUS0407141
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001509
; CURRENT APPLICATION NUMBER: PCT/US04/07141
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 44201
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 810
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-07141-810

Query Match      34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy      7  RGWLWALLRLRLAQSEPTTTHQPGYCAPYDECG-----KNPELSGLMTLSNVSCLS 59
Db      4  RGLALGLLILLCPAQVFSQ-----SCVMYGEIGAYGDKRYNCEYSG-----46

Qy      60  NTPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLRSLSTIKALLTRCPACSDN 119
Db      47  --PPKPLPKDGVLDVQELCPGFFFG-NVSLCCDVRQLQTLKDNQLQPLQFSRCSFCYN 103

Qy      120  FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYBAFYQHSFAEQSYDSCSRV 175
Db      104  LNLNLCFELTCSRQSOFLNVTATEDYDVPVNTQKTNVKELYYQGSFANAMYNACRDV 163

Qy      176  RVPAAATLAVGTMCYGYGALCNAQRWLNFGQDTGNGLAPLDIT-----PHLEPGQA 228
Db      176  RVPAAATLAVGTMCYGYGALCNAQRWLNFGQDTGNGLAPLDIT-----PHLEPGQA 228
```

Db 164 EAPSSNDKALGLLCKGKADA-CNATNWIEMFNKONGQAPFTITPVSPDPVH----- 215

Qy 229 VSGIOPLNEGVARCNESQGDVATCSCQCAASCPAIAIPQ-----ALDSTFYLG 279

Db 216 ---GMEPMNATKGCDSVDEVTAPCSCQDCSIVCGPKQPQPPPPAPPTWTLGLDAMVIM 272

Qy 280 QMPGSLVLIILCSFPAVVTILL-----VGFVAPARDKSKMVDPKGTLSLS 326

Db 273 WITTMAPLLVFFGAFPAVWCYRKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSVIPVVALAAGLVFTLTTPDVLWSAP 386

Db 325 DPVSAAPFEGCLRLFTRMGSCFVRNPGCVIPFSLVPIACSSGLVFRVTNTPVDLWSAP 384

Qy 387 NSQARSEKAFHDQHGFPFRNTQVILTAPNRSRYVDSLLLGPK-NFSGILDLDLLELL 445

Db 385 SSOARLEKEYFDQHGFPFRTEQLIIRAPLTDKHIYQYPSPGADVPFGPPLDIQILHVL 444

Qy 446 ELQERLRHLQVWSPQAQRNLSLODICVAPLNPNTSLYDCCINSLLQYFQNNRTILLTA 505

Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSYNT---NCTILSVLNYPQNSHVLHKK 499

Qy 506 NOTLMGQTSQVMDKDHFLYCANAPLTFKDGTTALALSCMADYGAPVFPFLAIGYKGYK 565

Db 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLHDPCLGTGFGVPFWLVLGDDYDQNYN 556

Qy 566 EBAALIMTFSLNYPAGDPRLAQAKLWEBAFLIEMRAFORRMAGMFQVTTAERSLEDEI 625

Db 557 NATALVITPPVNNYNDTEKLQRAQAWEKEFINFVKNYKN---PNLTISFTAERSIEDEL 613

Qy 626 NRTTADLPIFATSYIVIFLYISLALGSYSSRWVDSKATGLGCVAVVLGVANWAMG 685

Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRLLVDSKVSIGIAGILVLSVACSLG 673

Qy 686 FFSYLGRSLVILQVVPPLVLSVGADNIFIFVLEYQRLPRRGEPREVHIGBALGRVAP 745

Db 674 VFSYIGLPLTVIEVLPFLVLAGVDNIFILVQAYQORLQGETLQDLQGLRVLGEVAP 733

Qy 746 SMLLCSLSBAICFFLGALTPMAVRTFALTSLGLAVILDLLQMSAFVALLSDSKRQEAS 805

Db 734 SMLFSSFSFETVAFFLGALSMPAVHTFSLFAGLAVFIDFLQITCFVSLGLLDIKQEK 793

Qy 806 RLDVCCVQPOELPPGQ-GEGLLLGFQKAVAPFLLHMITRGVULLLFLALFGVSLYSM 864

Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLLDKWMRPIVIAIFVGLVSFSAVL 853

Qy 865 CHISVGLDQELALPKDSYLDLDFLFLNRYFEVGAUVYVTTLGYNFSSEAGNAICSSAG 924

Db 854 NKVDIGLQSLGMPDDSYMVDYFYSISQYLHAGPPVYFVLEBHDYTSKGQNMVCGMG 913

Qy 925 CNMFSTQKIQVATEPEQSYLAIPASSWDDPIDMLTP-SSCCRLYISGPNKDKFCPST 983

Db 914 CNNDLSVQIIFNAQLDNYTRIGFAPSSIDDYFDWVKPQSSCRV---DNITDQFCNAS 970

Qy 984 VNSLNLCKNMSIT-MGSVRPSVEQHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040

Db 971 VVDPACVR-CRPLTPREGKORPOGGDFMRFLPMLFSDNPNPKCKGKGHAAYSSAVNLLGH 1029

Qy 1041 DQOVLASRWMAHKLKNSQDYTEALRAARELANITADLRKVPGTDPAPFPPYITIVY 1100

Db 1030 GTRVGATYPMYVHTVLOTADFDIALKKARLTASNT-ETWGINGSG--AYRVPYVSFVY 1086

Qy 1101 FVEQYLTILPEGLFMLSCLPVTFAVSCLLGLDLRSGLNLLSIWMLVDVTGFMALWD 1160

Db 1087 FVEQYLTITDDYIFNLGVSLGAFLVTWVLGCELWSAVIMCATIAMLVNNFMGVNWLWG 1146

Qy 1161 ISYNAVSLINLAVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTN 1220

Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVRVERAREALAHMGSVFSGITLTK 1206

Qy 1221 LFGILVLGAKAQLQIQIFFRNLNLLITLGLHLGVLFPVLISYVGPDVNPA 1272

Db 1207 FGGIVVLAPAKSQIQIFYFRMYLAMVLIGATHGLIFLPLVLSYIGPSVKA 1258

RESULT 58

PCT-US05-18850-1032
; Sequence 1032, Application PC/TUS0518850
; GENERAL INFORMATION:
; APPLICANT: CHIN, Lynda
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: DPN-059-2
; CURRENT APPLICATION NUMBER: PCT/US05/18850
; CURRENT FILING DATE: 2005-05-27
; NUMBER OF SEQ ID NOS: 1528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1032
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US05-18850-1032

Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLAQSEPYTHIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59

Db 4 RGLALGLLGLLLCPAQVFSQ-----SCVMYGBGCIAYGDKRYNCEYSG----- 46

Qy 60 NTPARKITGDHLLILQKICPRLYTGNTQACSAKQOLVSEBASLITKALLTRCPACSDN 119

Db 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKONLQLPLOFLSRCPSCFYN 103

Qy 120 FVNLHCNTCSNOSLFINVTR-----VAQIGAGQLPAVVAYBAFYOHSPABOSYDSCSRV 175

Db 104 LNLFPCLTCSRQSOFLNVTATEDYVDPVNTQNTNKKELQYVYGQSPANAMYNACRDV 163

Qy 176 RVPAATLAVGTCMGYVGSALCNAQRLNFGQDGTGNGLAPLDIT-----PHLLSPGQA 228

Db 164 EAPSSNDKALGLLCKGKADA-CNATNWIEMFNKONGQAPFTITPVSPDPVH----- 215

Qy 229 VSGIOPLNEGVARCNESQGDVATCSCQCAASCPAIAIPQ-----ALDSTFYLG 279

Db 216 ---GMEPMNATKGCDSVDEVTAPCSCQDCSIVCGPKQPQPPPPAPPTWTLGLDAMVIM 272

Qy 280 QMPGSLVLIILCSFPAVVTILL-----VGFVAPARDKSKMVDPKGTLSLS 326

Db 273 WITTMAPLLVFFGAFPAVWCYRKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSVIPVVALAAGLVFTLTTPDVLWSAP 386

Db 325 DPVSAAPFEGCLRLFTRMGSCFVRNPGCVIPFSLVPIACSSGLVFRVTNTPVDLWSAP 384

Qy 387 NSQARSEKAFHDQHGFPFRNTQVILTAPNRSRYVDSLLLGPK-NFSGILDLDLLELL 445

Db 385 SSOARLEKEYFDQHGFPFRTEQLIIRAPLTDKHIYQYPSPGADVPFGPPLDIQILHVL 444

Qy 446 ELQERLRHLQVWSPQAQRNLSLODICVAPLNPNTSLYDCCINSLLQYFQNNRTILLTA 505

Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSYNT---NCTILSVLNYPQNSHVLHKK 499

Qy 506 NOTLMGQTSQVMDKDHFLYCANAPLTFKDGTTALALSCMADYGAPVFPFLAIGYKGYK 565

Db 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLHDPCLGTGFGVPFWLVLGDDYDQNYN 556

Qy 566 EBAALIMTFSLNYPAGDPRLAQAKLWEBAFLIEMRAFORRMAGMFQVTTAERSLEDEI 625

Db 557 NATALVITPPVNNYNDTEKLQRAQAWEKEFINFVKNYKN---PNLTISFTAERSIEDEL 613

Qy 626 NRTTADLPIFATSYIVIFLYISLALGSYSSRWVDSKATGLGCVAVVLGVANWAMG 685

Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRLLVDSKVSIGIAGILVLSVACSLG 673

QY 1221 LPGILVILGAKAQLIOIFRFRNLTLILGLHLGLVFLPVILSYGPDVNP 1272
Db 1207 FGIIVVLAFAKSIQIFIRFRMYLAMVLLGATHGLIFLVLISYIGPSV 1258

RESULT 60
US-10-208-731-2
; Sequence 2, Application US/10208731
; GENERAL INFORMATION:
; APPLICANT: Carstean, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/10/208,731
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-208-731-2

Query Match 34.8%; Score 2402.5; DB 32; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGMILWALLRLAQSEPTTIHQPGYCAFYDECG-----KNPELSSGLMTLSNVCSLS 59
Db 4 RGLALGLLLLLCPAQVFSQ-----SCWYGECCGAYGDKRYNCYSG-----46

QY 60 NTPARKITGDHLLILQKICPRLYTGENTQACSAKQLVLSLEASLITKALTRCPACSDN 119
Db 47 --PPKPLPKGDYDLVQELCPGFFFG-NVSLCCDVRLQTLKDLQLPLQFLSRCSFCFN 103

QY 120 FVNLCHTNCSPQSLFNVT-----VAQLGAGQIPAVVAYEAFVQHGFABOSYSCSRV 175
Db 104 LNLFLCELTCSPROQFLNVTATEDYDPVTNTQKTNKELQYVYVQGFANAMYNACRDV 163

QY 176 RVPAAATLAVGTCMGVYGSALCNAQRLNPFQDGTGNGLAPLDIT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEYMFNKDNGQAPFTTTPVFSDFPVH-----215

QY 229 VGSIGIQLNEGVARNESQDDVATCSQDCQCAASCPATARPQ-----ALDSTFYLG 279
Db 216 ---GMEPNWNAKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272

QY 280 QMPSGLVLIILCSFVAVTILL-----VGFVRVAPARDKSMVDPKKGTSL 326
Db 273 WITYMAFLVFGAFVAVCYKRYFSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

QY 327 DKLSFSTHLLQFQCGWGTWASWPLTILVSLVPVVALAAGLVFTLTTPDVLWAP 386
Db 325 DPVSAFEGCLRLFRWGSCFVRPGVGVIFPSLVFITAACSGLVFVRVTNPNVDLWAP 384

QY 387 NSQARSEKAFDHOHGPFRFTNOVTLTAPNRSSRYSDLSLLGPK-NFSGILDLDLLELL 445
Db 385 SSQARLEKEYFDHGPFRFTNOVTLTAPNRSSRYSDLSLLGPK-NFSGILDLDLLELL 444

QY 446 ELQERLRHLQWSPQAQRNLSQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASYDNETVTLDICLAPLSYNI---NCTILSVLYNFQNSHVLHKK 499

QY 506 NOTLMGQTSQVDWKDHFYLCANAPITFKDGTALALSCMADYGAVPFPFLAIGGYKDKYS 565
Db 500 GDDEF--VIADYHTHFLCYRAPASLNDTSLHDPCLGTGPGVPFVPLVGLGDDQNY 556

QY 566 EBAALIMTFSLNYPAGDPRRLAQAKLWEAPLEEMRAFORRMAGMPQVTTTARSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKNYKN---PNLTISFTARSSTED 613

QY 626 NETTAEDLPATSYIVFLYISLALGSSYSSSRVWVDSKATILGLGGVAVLGAAMG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDSKVSIGIAGILVILSSVACSLG 673

QY 686 FFSYLGRSSVLQVVPFLVSGADNIFIFVLEYQRLPRRPGEPREHIGRACRVAP 745
Db 674 VFSYIGLPLTLIVIEIPFLVAVGVDNIFILVQAYQORDERLQGETLDQOLGRVLEAP 733

QY 746 SMLCSSLSEACFFLGAFTMPAVRTFALTSGVLVLDLQWAFVALLSDSKQREAS 805
Db 734 SMFLSSFSETVAFFLGALSVMVAVHTFSLFAGLAVFIDELLQITCFVSLGLDQKQKN 793

QY 806 RLDVCCCKVQELPPGQ-GEGLILGFPQKAYAPELLHWTIGVLLLLFLALFGVLSYM 864
Db 794 RLDIFCCVRGABEDGTSVQASESCLFRFKNSYSPILLKDDMRPIVIAIFVGLSFSIAVL 853

QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVYFVTLTGYNFSSSEAGNAICSSAG 924
Db 854 NKVDIGLDQSLSDMPDDSYWVDYFKISQVYLHAGPPVYFVLEEGHDTSSKGQNMVCGMG 913

QY 925 CNMFSTQKIQYATEPPEQSYLAIPASSWVDDFIDMLTP-SSCCRLYISGPNKDKFCPST 983
Db 914 CNMDSLVOQIFNAALDNYTRIGPAPSSWIDDYDFWVKPQSSCCRV---DNITDQFCNAS 970

QY 984 VNSLNLKNCMSIT-MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLLAAYSIVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNKPCGKGHAAYSSAVNLLGH 1029

QY 1041 DQOVLSRPMVHKPKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVATYFMYHTVLTQISADFIDALKKARLIASNT-ETMGINGS--AIRVFPYSVTV 1086

QY 1101 FYEQYTLTPEGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLIVDTVGFMAWD 1160
Db 1087 FYEQYTLTIIDTIFNLGVSGLGAIPLVTVLGLCELWSAVIMCATIAMLVNMGVWMLWG 1146

QY 1161 ISYNAVSLINLAVGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVPAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGIGSVEFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK 1206

QY 1221 LPGILVILGAKAQLIOIFRFRNLTLILGLHLGLVFLPVILSYGPDVNP 1272
Db 1207 FGIIVVLAFAKSIQIFIRFRMYLAMVLLGATHGLIFLVLISYIGPSV 1258

RESULT 61
US-10-218-140-5494
; Sequence 5494, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218,140
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 5494
; LENGTH: 1278
; TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-218-140-5494

Query Match      34.8%; Score 2402.5; DB 32; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLWALLLLRLAQSEPTTHHQPCYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECCGAYGDKRYNCBSYG----- 46
QY 60 NTPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKDNLQLPLOFLSRCPSCFYN 103
QY 120 FVNHLCHNTCPNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAESQSDSCSRV 175
DB 104 LLNLFCELTCSPRQSOFLNVTATEDYDVPVNTQTKNVKELQYYVVGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCVGYGSALCNAQRWLNFGQDTGNGLAPLDIT-----PHLLEPGQA 228
DB 164 EAPSSNDKALGLLCCGDADA-CNATNWIEYMFNKONGQAPFTITPVFSDFPVH----- 215
QY 229 VGSQIQLNEGVARCNESQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTLGLDAMVIM 272
QY 280 QMPGSLVLIILLCVSFVAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLLVFGAFVAVCMYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DXLSSTHTLLQFQCGWGTWASWPLTILVLSVTPVVALAAGLVFETLTDVPELWSAP 386
DB 325 DPVSAAFEGCLRLTRMGSCFCVRNPGCVIFFSLFITACSSGLVFRVTTNPVDLWSAP 384

RESULT 62
US-10-741-600-1542
; Sequence 1542, Application US/10741600
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1542
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1542

Query Match      34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLWALLLLRLAQSEPTTHHQPCYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECCGAYGDKRYNCBSYG----- 46
QY 60 NTPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKDNLQLPLOFLSRCPSCFYN 103
QY 120 FVNHLCHNTCPNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAESQSDSCSRV 175
DB 104 LLNLFCELTCSPRQSOFLNVTATEDYDVPVNTQTKNVKELQYYVVGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCVGYGSALCNAQRWLNFGQDTGNGLAPLDIT-----PHLLEPGQA 228
DB 164 EAPSSNDKALGLLCCGDADA-CNATNWIEYMFNKONGQAPFTITPVFSDFPVH----- 215
QY 229 VGSQIQLNEGVARCNESQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTLGLDAMVIM 272
QY 280 QMPGSLVLIILLCVSFVAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLLVFGAFVAVCMYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DXLSSTHTLLQFQCGWGTWASWPLTILVLSVTPVVALAAGLVFETLTDVPELWSAP 386
DB 325 DPVSAAFEGCLRLTRMGSCFCVRNPGCVIFFSLFITACSSGLVFRVTTNPVDLWSAP 384
```

```
QY 387 NSQARSEKAFHDPHQFPGFRTNQVILTAPENRSSRYRDSLLLPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKEBYDFHQFPGFRTNQVILTAPENRSSRYRDSLLLPK-NFSGILDLDLLELL 444
QY 446 ELQERLRHLQVMSPEAQRNLSLQDICYAPLNPNTSLYDCCINSLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASYDNETVTLQDCLAPLSPYNT--NCTILSVLNYFQNSHVLDDHK 499
QY 506 NOTLMGQTSQVMDKDFLYCANAPLTFKOGTALALSCMADYDAPVPFPFLAIGYKGDYS 565
Db 500 GDDFF--VYADYTHFLYCVRAPASLNDTSLHDPCLGTFGGVPFPLVLGGYDQNYN 556
QY 566 EAEALMTFSLNYPAGDPRLAQAKLWEAEFLMEAFQRMAGFQVTFPFAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLOQAQWKEFNFVKNYKN--PNLTISFTASRDEDEL 613
QY 626 NRTTAEPLIFATSYIVFIYLSALGSYSRSMRVMDSKATLGLGVAVVLGAVMAAMG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRELLVDSKVSIGIAGILVLSSVACSLG 673
QY 686 FFSYLGRSLVILQVVPFVLAVSGADNIFIVLEYQRLPRRPGEPREVIHGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFPLVAVGVNDIFILVQAYQORDERLQGETLDQOLGRVLEVP 733
QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSDSKREAS 805
Db 734 SMFLSSFSSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLDKEQKN 793
QY 806 RLDVCCVQKQELPPPGQ-GEGLLGFQKAYAPFLHMTIRGVVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKDWMRPIVIAIFVGLSFI AVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVFTVTLGNYFSSBAGMNAICSSAG 924
Db 854 NKVDIGLQDLSNPDSDYMDYFKSISQYLHAGPPVYFVLEEGHDTYSSKGQNMVCGMG 913
QY 925 CNNFSTQKIYATEPPEOSYLAIPASSWDDFDLWLP--SSCCRLYISGPNKDKPCPST 983
Db 914 CNNDLSVQOIFNAALQDNYTRIGFAPSSWIDDYFDMVKPQSSCCRV--DNITDQPCNAS 970
QY 984 VNSLNLCKNCSMT-MGSVRPSVEQFKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR-CRPLTPGKORPQGGDFMFLPMLSDPNPNPKCGKGHAAYSSAVNTLLGH 1029
QY 1041 DGQVLASRFMAHYHKLKNSQDYTEALRAARELAANTADLRKVPYGTDPFAFVPPYTTIN 1100
Db 1030 GTRVGATYFTWYTHVLOTGADFDALKARLIASNT-ETMGINGS--ATRVFPYSVTV 1086
QY 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLISYIMVLVDTVGFMAWD 1160
Db 1087 FVEQYLTIIDDTIFNLGVSIGALFVTLVGLGCELSAVIMCATIAMVLVNMFGVMWLMG 1146
QY 1161 ISYNAVSLNLVSAVMSVEFVSHITRPSAISTKPTWLERAKATISMGSAVPAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSVCSISVEFCSHITRAFTVSMKGSRVERAEBALAHMGSSVFGITLTK 1206
QY 1221 LRGILVLGAKAQLOIOPFRNLNLTLLGLHGLVFLPVLISYVGPVNPA 1272
Db 1207 FGVIVVLAFAKSOIQIFRMYLAVMLVGLATHGLIFLPLVLSYIGPSVNKA 1258
```

RESULT 63

```
US-10-741-601-530
; Sequence 530, Application US/10741601
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 530
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-530
```

```
Query Match 34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGWLLWALLRLAQAQSEPTTIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
Db 4 RGLAUGLLLLLLCPAFVFSQ-----SCVMYGEGIAYGDKRYNCRYS-- 46
QY 60 NTPARKITGDHLLILQKICPRLYTGNTQACSAKQVLSKASLSTTKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVDELCPGFFG--NVSLLCCDVRLQTLKONLQPLQFLSKPCSPCFYN 103
QY 120 FVNLHCHNTCSNQSLFINVTR---VAQIGAGQLPAVVAEAFYQHSFAESQYDSCSRV 175
Db 104 LLNLFCELTCSQRSQFLNVTATEDYVDPVTNQTKTNKVELQYVVGQSPANAMYNAACRDV 163
QY 176 RUPAAATLAVGTWCGYVGSALCNAQRWLNFOGDTGNGLAPLDIT-----FHLEPGQA 228
Db 164 EAPSSNDKALGLLGGKADADA--CNATWIBYMFKNKGQAPFTITPVFSDPPVH----- 215
QY 229 VSGGIQPLNEGVARCNESQGDVATCSCODCAASCAPAIARPQ-----ALDSTFYL 279
Db 216 ---GMEPMNATKGCDSDEVTAPCQDCSIVCGPKQPPPPPPAPWTILGDAMVIM 272
QY 280 QMPSGLVLIILCSVFAVVTILL-----VGRFVAPAPARKSKMVDPKGTSL 326
Db 273 WITWMAFLVFGAPFAVFCYKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLQOPFGNGTWTWASWPLTLVLVSVIPVVALAGLVTELTDDPVELMSAP 386
Db 325 DFVSAFEGCLRLRLETRWGSFCVRNPGCVIFFSLVFITACSSGLVFRVTVNTEVDLMSAP 384
QY 387 NSQARSEKAFHDPHQFPGFRTNQVILTAPENRSSRYRDSLLLPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKEBYDFHQFPGFRTNQVILTAPENRSSRYRDSLLLPK-NFSGILDLDLLELL 444
QY 446 ELQERLRHLQVMSPEAQRNLSLQDICYAPLNPNTSLYDCCINSLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASYDNETVTLQDCLAPLSPYNT--NCTILSVLNYFQNSHVLDDHK 499
QY 506 NOTLMGQTSQVMDKDFLYCANAPLTFKOGTALALSCMADYDAPVPFPFLAIGYKGDYS 565
Db 500 GDDFF--VYADYTHFLYCVRAPASLNDTSLHDPCLGTFGGVPFPLVLGGYDQNYN 556
QY 566 EAEALMTFSLNYPAGDPRLAQAKLWEAEFLMEAFQRMAGFQVTFPFAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLOQAQWKEFNFVKNYKN--PNLTISFTASRDEDEL 613
QY 626 NRTTAEPLIFATSYIVFIYLSALGSYSRSMRVMDSKATLGLGVAVVLGAVMAAMG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRELLVDSKVSIGIAGILVLSSVACSLG 673
QY 686 FFSYLGRSLVILQVVPFVLAVSGADNIFIVLEYQRLPRRPGEPREVIHGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFPLVAVGVNDIFILVQAYQORDERLQGETLDQOLGRVLEVP 733
QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSDSKREAS 805
Db 734 SMFLSSFSSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLDKEQKN 793
QY 806 RLDVCCVQKQELPPPGQ-GEGLLGFQKAYAPFLHMTIRGVVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLKDWMRPIVIAIFVGLSFI AVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVFTVTLGNYFSSBAGMNAICSSAG 924
Db 854 NKVDIGLQDLSNPDSDYMDYFKSISQYLHAGPPVYFVLEEGHDTYSSKGQNMVCGMG 913
QY 925 CNNFSTQKIYATEPPEOSYLAIPASSWDDFDLWLP--SSCCRLYISGPNKDKPCPST 983
Db 914 CNNDLSVQOIFNAALQDNYTRIGFAPSSWIDDYFDMVKPQSSCCRV--DNITDQPCNAS 970
QY 984 VNSLNLCKNCSMT-MGSVRPSVEQFKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR-CRPLTPGKORPQGGDFMFLPMLSDPNPNPKCGKGHAAYSSAVNTLLGH 1029
QY 1041 DGQVLASRFMAHYHKLKNSQDYTEALRAARELAANTADLRKVPYGTDPFAFVPPYTTIN 1100
Db 1030 GTRVGATYFTWYTHVLOTGADFDALKARLIASNT-ETMGINGS--ATRVFPYSVTV 1086
QY 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLISYIMVLVDTVGFMAWD 1160
Db 1087 FVEQYLTIIDDTIFNLGVSIGALFVTLVGLGCELSAVIMCATIAMVLVNMFGVMWLMG 1146
QY 1161 ISYNAVSLNLVSAVMSVEFVSHITRPSAISTKPTWLERAKATISMGSAVPAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSVCSISVEFCSHITRAFTVSMKGSRVERAEBALAHMGSSVFGITLTK 1206
QY 1221 LRGILVLGAKAQLOIOPFRNLNLTLLGLHGLVFLPVLISYVGPVNPA 1272
Db 1207 FGVIVVLAFAKSOIQIFRMYLAVMLVGLATHGLIFLPLVLSYIGPSVNKA 1258
```



```
QY 925 CNNSFTOKIQVATEFPEQSYLAIPASSWVDDFIDWLTP--SSCCRLYISGPNKQKFCPST 983
DB 914 CNNSLVQOIFNAAQLDNYTRIGFAPSSWIDDYFDWVRPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNCNKCMSIT--MGSVRPSVEQFHKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
DB 971 VDPACVR--CRPLTPGKQRPQGGDMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQVILASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNV 1100
DB 1030 GTRVGATYFMYHTVTLQTSADFDALKKARLIASNT--ETMGINGS--AYRVFPYSVFYV 1086
QY 1101 FVEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGPMALWD 1160
DB 1087 FVEQYLTIIDDTIFNLGSLGAIFLVTWVLLGCELWSAVIMCATIAMVLNMGVWMLWG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFAISKTWLERAKEATISMGSAVAGVAMTN 1220
DB 1147 ISLNAVSLNLSVCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
QY 1221 LFGIIVLGLAKAQLQIIFFRNLNLLITLGLHLGLVFLPVILSYVGPDPVNP 1272
DB 1207 FGGIIVLAFAKSQIQIFFRMYLAMVLLGATHGLIFLPLVLSYIGPSVYKA 1258

RESULT 64
US-10-756-149-4924
; Sequence 4924, Application US/10756149
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4924
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4924

Query Match 34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLMALLRLAQSEPTTIHQPGYCAFYDSCG-----KNPELSSGLMTLSNVCLSL 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCWYGGCGIAYGDKRYNCEVSG----- 46
QY 60 NTPARKITGDHILILQKICPRLYTGPNTQACCSAKQLVSLASLSITTKALLTRCPACSDN 119
DB 47 --PPRLPKDGYDLVQELCPGFFG--NVSLCCDVRLQTLKNLQLPLQFLSRCSFCFN 103
QY 120 FVNLHCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAESYDSCSRV 175
DB 104 LNLFLCELTCSPQSQFLNVTATEDYDVPVNTQTKNVKELQYVYVQGSFANAMYNACRDV 163
QY 176 RYPAATLAVGTMCYVGSALCNAQRLNFGQDTGNGLAPLDT-----PHLLEPGA 228
DB 164 EAPSNNDKALGLLCKGDADA--CNATNWIEMFNKONGQAPFTTIPVFSDFPVH----- 215
QY 229 VGSQIPLNEGVARCNESQGDVATCSQDCQCAASCPAIARPO-----ALDSTFYLG 279
DB 216 ---GHEPMNNAKGCDESDEVTAPCSQDCSIVCGPKQPPPPAPWILGLDAMYVIM 272
QY 280 QMPGSLVLIILCSFPAVVITLL-----VGRFVAPARDKSKMDPKGKTSLS 326
DB 273 WITYMAFLLVFGAFPAVVCYRKRYFVSBYTTIDSNIAFSV--NASDKGE-----ASCC 324
```

```
QY 327 DKLSFSTHTLLGQFPQGGTWTWASWPLTILVLSVIPVVALAAGLVFTBELTDPVLSWAP 386
DB 325 DPVSAAPGECRLRLFTRWGSCFVCPGCVIFESLVTITACSSGLVFRVTINPVDLWSAP 384
QY 387 NSQARSEKAFHDQHFPGFPFRNQVILTPAPNSSVRYDSLLILGPK--NPSGILDLDDLLELL 445
DB 385 SSQARLEKEYDQHFPGFPFRTEQLIIRAPLTKHIIQYPYSGADVPFGPPLDIQLHQLV 444
QY 446 ELQERLRLHQLWSPEAQNRNISLQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLDICLAPLSPYNT---NCTILSVINTFQNSHLDHKK 499
QY 506 NQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALALSCMADYCAPFPFLAIGYKGYKYS 565
DB 500 GDDFP--VYADYTHFLYCYRAPASLNDTSLHDPCLGTGPGVPFWLVLGGYDQNYN 556
QY 566 EBAALIMTFSNNYPAGDPRLAQAQKLWBEAELEENRAFQRMAGMFWQVTFTHAERSLEDBI 625
DB 557 NATALVITFPVNNYNDTEKLQRAQWKEKTFNFVKNYKN---PNLTISFTAERSTEDBL 613
QY 626 NRTTAEDLPFATSIVIVFLYISLALGSYSWSRVMVDSKATLGLGGVAVVLGAYMAAG 685
DB 614 NRESDSVFTVVISYAIMFLYISLALGHIKSCRRLLVDSKVSGLAGILIVLUSSVACSUG 673
QY 686 FFSYLGIRSSLVILQVVPFLVLSGADNIFIFVLEYQLPRRPGEPREPHIGHGRALGRVAP 745
DB 674 VFSYIGLPLTIVIEVIFVLAVGVNDIFILVQAYQDERLQGETLQOGLGRVLGEVAP 733
QY 746 SMLJCSLSEACFPFLGALTMPAVRTPALTGLAVILDPLQMSAFVALLSDSKRQES 805
DB 734 SMFLSSFSETVAFFLGALSVMPPAVHTFSLFAGLAVFIDFLQLITCFVSLGLDIKRQEN 793
QY 806 RLDVCCCKVPQELPPPGQ--GEGLLGFFQKAYAPPELLHWITRGVVLPLALFGVLSYM 864
DB 794 RLDIFCCVCRGAEADGTSVQASECLFRFPKNSVSPLLKDMWRPIVIALFVGLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLPLNRYFEVGAPVYFVTTLGYNFSSEAGNNAICSSAG 924
DB 854 NKVDIGLQSLSPDDSYWDYFKSISOYLHAGPPVYFVLEBHDYTSKQGNMVGCGMG 913
QY 925 CNNSFTOKIQVATEFPEQSYLAIPASSWVDDFIDWLTP--SSCCRLYISGPNKQKFCPST 983
DB 914 CNNSLVQOIFNAAQLDNYTRIGFAPSSWIDDYFDWVRPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNCNKCMSIT--MGSVRPSVEQFHKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
DB 971 VDPACVR--CRPLTPGKQRPQGGDMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQVILASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNV 1100
DB 1030 GTRVGATYFMYHTVTLQTSADFDALKKARLIASNT--ETMGINGS--AYRVFPYSVFYV 1086
QY 1101 FVEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSVIMILVDTVGPMALWD 1160
DB 1087 FVEQYLTIIDDTIFNLGSLGAIFLVTWVLLGCELWSAVIMCATIAMVLNMGVWMLWG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFAISKTWLERAKEATISMGSAVAGVAMTN 1220
DB 1147 ISLNAVSLNLSVCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
QY 1221 LFGIIVLGLAKAQLQIIFFRNLNLLITLGLHLGLVFLPVILSYVGPDPVNP 1272
DB 1207 FGGIIVLAFAKSQIQIFFRMYLAMVLLGATHGLIFLPLVLSYIGPSVYKA 1258
```

RESULT 65

US-10-796-280-1261

```
; Sequence 1261, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001510
```



```
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1261
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-796-280-1261

Query Match          34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGMWALLRLRLAQSEPYTHIQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DB 4 RGLALGLLGLLLCPAQVFSQ-----SCVMYGEGLAYGDKRYNCEVSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGVLDVQELCPGFFG-NVSLCCDVRQLQTLKDNQLPLQFLSRCPSCFYN 103

QY 120 FVNLCNTCSNQSFLINVTNTR-----VAQLGAGQLPAAVVAEAFYQHSFAGSYPSCSRV 175
DB 104 LNLNLCBELTCSRQSOFLNVTATEDYDVPVNTQNTKYNKELQYVVGQSPANAMYNACRDV 163

QY 176 RVPAAATLAVGTMCVGYGSCALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLEPGA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVFSDFPVH-----215

QY 229 VSGIQLNEGVARNCSGQDDVATCSQDCAASCAPATARPQ-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYVIM 272

QY 280 QMPGSLVLIILLCSPFAVVITILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITTYMAFLVFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKG-----ASCC 324

QY 327 DKLSFTSTHLLQFFQGWGTWASWPLTILVLSVPVVALAAGLVFTLTTPDVELWSAP 386
DB 325 DPVSAAFEGCLRLTRWGSFCVRPGCVIFSLVPIFACSSGLVFRVTNTPVLDWSAP 384

QY 387 NSQARSEKAFDHFQPPRTNQVILTAPNRSRYDLSLLGPK-NFSGILDLDLLELL 445
DB 385 SSOARLEKEFYDFHFGPPRTQLITRAPLTDKHIYQYPVPGADVPFGPPLDIQLHQL 444

QY 446 ELQERLRHLQWSPENQRNLSIDICYPALNPONTSLYDCCINSILLOYFQNNRTLLITA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSYNT---NCTILSVLNYFQNSHVLHKK 499

QY 506 NQTLMGQTSQVDKDHFLYCANAPLTFKDGTAALSCMADYGAAPVFPFLAIGYKGDYS 565
DB 500 GDDFPF---VYADYHFLYCVAPASLNDTSLHBPCLGTGFGPFPVPLVGLGYDDQYN 556

QY 566 EBAALIMTFSLNYPAGDPRLAQAKLEBAFLEEMRAFORRMAGMFQVTFTAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKLQRAQAEKEFINFKVKNYKN---PNLTISFTAERSIEDEL 613

QY 626 NRTTAEPLIFATSVIVIFLYLSLALGYSYSSNRWVDSKATLGGVAVVLGAVNAGM 685
DB 614 NRESDDVFTVVISVAINMFLYLSLALGHITKSCRRLLVDSKVSLSIAGILVLSVACSLG 673

QY 686 PFSYLGRSSVLQVWPPLVLSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
DB 674 VESYIGLPLTLVIEVIFPLVAVGVNDIFILVQAYQDERLOGETLDOOLGERVLEAVP 733

QY 746 SMLLCSLSBAICFFLGAFTMPAVRTFALTSGVLAVLDLFLQMSAFVALLSDSKRQES 805
DB 734 SMFLSSFSBETVAFGLSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIKQEK 793

QY 806 RLDDVCCVKPQLPPLPGQ-GEGLLIGFQKAYAPFLHWTIRGVVLLFLALFGVLSYM 864
DB 794 RLIDIFCCVRGAEDGTSSQASESCLFRFFKNYSPLLLKDMWRPIVIAIPVGLVLSFIAVL 853
```

RESULT 66

```
US-10-796-307-810
; Sequence 810, Application US/10796307
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001509
; CURRENT APPLICATION NUMBER: US/10/796,307
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 44201
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 810
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-796-307-810
```

```
Query Match          34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGMWALLRLRLAQSEPYTHIQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DB 4 RGLALGLLGLLLCPAQVFSQ-----SCVMYGEGLAYGDKRYNCEVSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGVLDVQELCPGFFG-NVSLCCDVRQLQTLKDNQLPLQFLSRCPSCFYN 103

QY 120 FVNLCNTCSNQSFLINVTNTR-----VAQLGAGQLPAAVVAEAFYQHSFAGSYPSCSRV 175
DB 104 LNLNLCBELTCSRQSOFLNVTATEDYDVPVNTQNTKYNKELQYVVGQSPANAMYNACRDV 163

QY 176 RVPAAATLAVGTMCVGYGSCALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLEPGA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVFSDFPVH-----215

QY 229 VSGIQLNEGVARNCSGQDDVATCSQDCAASCAPATARPQ-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYVIM 272

QY 280 QMPGSLVLIILLCSPFAVVITILL-----VGRVAPARDKSKMVDPKKGTSL 326
```

Db	273	WITWMAELLVFFGGAFFAWCVRKRYFVSEYTPIDSNIAFSV-NASDRGE-----ASCC	324
Qy	327	DKLSFSHTLLGQFFQCGTGWASWPLTILVLSVIPVALAAGLVFTELTTDPVELWSAP	386
Db	325	DPVSAAFEGCLRLLFTRWGSFCVRNPGCVIFFSLVETACSSGLVFRVVTNPNVPLWSAP	384
Qy	387	NSQARSEKAFHDQHFGFPFRTNQVILTAPNRSSRYDSLLGPK-NPFGILDLDLLELL	445
Db	385	SSQARKEKEYFDQHFGFPFRTQVILRAPLTDKHIIQYPFSGADVPFGPPLDIQILHQVL	444
Qy	446	ELQERLHLQVWSEPAQRNLSLODICVAPLNPDTNLSLYDCINSLLOYFONNRTLLLLTA	505
Db	445	DLQIAIEN--ITASYDNETVLODLCIAPISPYNT---NCTILSVLNYFNQSHSVLDHKK	499
Qy	506	NOTLMGOTSQVDMKOHFLYCANAPLTPKDGTAALSCMADYGAPVPFPFLAIGYKGDYS	565
Db	500	GDDPFF--VYADYHTHELYCVRAPASLNDATSLLDHPCGLTGGFPVFPMLVGLGYDQNYN	556
Qy	566	EAEALIMTFSLNPNYPAGDPRLAQAKEEAFLEEMRAFORRMAGMFOVTTAERSLEDEI	625
Db	557	NATALVITFPFNYYNDETEKLQRAQAEKGFNFVKNYKN---PNLTISTAGRSIEDLE	613
Qy	626	NRTAEDLPFATSYIYIFLYISLALGSYSWSRWVMDSKATILGLGVAIVVLGAVMAAG	685
Db	614	NRESDSVFTWISYAIMFLYISLALGHKISCRLLVDSKSLGIAGILVLSVACSLG	673
Qy	686	PFSVLGTRSSILVLOVFPVLVSVGADNIIFVLRYQRLPRRGPPEPREVHIGRALGRVAP	745
Db	674	VFSYIGLPLTLVIEVLPFLVAVGVNITFILVQAYQORDERLOGETLDQGLGRVLGEVAP	733
Qy	746	SMLLCSLSEAICPFGLALTMPAPVTRPALTSGLAVIDLFLQMSAFVALLSLDSKQKEAS	805
Db	734	SMFLSSFSFETAVFLGALSVMPAVHTSFAGLAVFIDFLQITCFVSLILGLDIKQEKV	793
Qy	806	RLDVCCCKVQKQELPPPGQ-GEGLLGFQKAYAPFLHLHITRGVLLFLFALFGVSLYSM	864
Db	794	RLDIFCCVRGAEDGTSVQASESLCFRPFKNYSPLLLKDMWRPVIATFVGVLSFSIAVL	853
Qy	865	CHTSVGLDQELALPKDSYLDYFLFNRYFEVGAPVYFVTTGLVNSSEAGMAKICSSAG	924
Db	854	NKVDIGLQDSLWPDSDSYWDYFKSISQYLHAGPPVYFLBEGHDYTSXGQNMVCGMG	913
Qy	925	CNNFSFTQIKYATEPPEQSYLAIPASSWDDDFIDMLTP--SSCCRLYISGFNKDKFCPST	983
Db	914	CNNDLSLVQIIFNAALQDNYTRIGAPSSWIDDYFDMVKPQSSCCRV--DNITDQFCNAS	970
Qy	984	VNSLNLCKNCKMSIT-MGSVPSVSEQPHKYLFWFLNDRPNIKCPKGGILAAYSTSVN--LTS	1040
Db	971	VVDPAQVR-CRPTPEBKQRPQGGDFWRFLPMLFSDNPNPKCGKGGHAAVSSAVNTLLGH	1029
Qy	1041	DGOVLSRRFMAYHKPLKNQOYTEALRAARELAANITADLRKVPGTDPAPVEVPYTTNV	1100
Db	1030	GTRVGATYFTYHTVLQTSADFIDALKARLLASNVT-ETWINGS--ATRVFPYSVPYV	1086
Qy	1101	FYEQYLTILPEGLPMLSCLVPTFVASCILLGLDRLSGLNLISIVMLIVDTVGVFNALWD	1160
Db	1087	FYEQYLTIIIDDTIFNLGVSGLAIFLVTMVLGCELWSAVTMCRATIAMVLVNMFGVMWLWG	1146
Qy	1161	ISYNAVSLINLVAAGMSVFEVSHITRSPAI STKPTWLERAKATTISMGSAVAGVAMTN	1220
Db	1147	ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKGSVERAREALAHMGSSVFGITLTK	1206
Qy	1221	LPGLVLGLAKAQIQTIFFFRNLNLTLLGLLHGLVFLPVLVSYVGDVNPVA	1272
Db	1207	FGGIVLAFAKSGTFOIIFYRMTIAMVLLGATGGLIPLPVLVSYGSPVNA	1258

RESULT 67

US-10-821-801-949

US-10-021-901-949
; Sequence 949, Application US/10821801

; GENERAL INFORMATION:

; APPLICANT: Hinzmann, Bernd

```

Db      614  NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDKSVLSIGIAGILVLSSVACSLG 673
Qy      686  FFSYLGRISLVILQVVPFLVLSVGADNIPFVLEYQRLPRRPGEPREHIGRALGRVAP 745
Db      674  VFSYIGLPLTLVIEVIFPLVLAAGVDNIFILVQAYORDERLQGETLDQOLGRVLEAVP 733
Qy      746  SMLLCSLSEAI CFFLGALTPMPAVRTPALTSGVLAVILDFLLQMSAFVALLSLDSKQREAS 805
Db      734  SMLFSSFSSTVAFFLGLSMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKRQKKN 793
Qy      806  RLDDVCCVKPQELPPPGQ-GEGLLGFQKAYAPFLLHMTIRGVLLFLFALFGVSLYSM 864
Db      794  RLDFCFCVGAEDGTSVQASESCLFRFFKNSYSPLLLKDMRPVIAIFVGLSFSIAVL 853
Qy      865  CHISVGLDQELALPKDSYLLDYFLNRYFEVGAFFVFTTLLGNFSSBAGMAICSSAG 924
Db      854  NKVDIGLQSLSPDDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKGGQNMVCGMG 913
Qy      925  CNPFSTQKIQVATEPPEQSYLAIPASSWDDPIDWLTP-SSCCRLYISGPNKDKFCPST 983
Db      914  CNNDSLVQOIFNAAQLDNRYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy      984  VNSLNLCKNCSIT-MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db      971  VDDPACVR-CRPLTPEGKORPOGGDFMRFLPMLPFLSDNPNPKCKGGAAYSSAVNTLLGH 1029
Qy      1041  DGQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFEPYPTITNV 1100
Db      1030  GTRVGATYFTVHTVLOTADFDALKKARLIASNT-ETMGINGS--AYRFPFVSFVY 1086
Qy      1101  FFEQYLTLPEGLFMLSCLVPTFAVSCLLLGLDLSGLLNLISIVMILVDTVGFMAIWD 1160
Db      1087  FFEQYLTIIDDTIFNLGVSIGALFVLTWVLGCELWSAVIMCATIAMVLVNMFGVMWLMG 1146
Qy      1161  ISNAVSLNLVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMSAPGAVMTN 1220
Db      1147  ISNAVSLNLVNSCGISVEFCSHITRAFTVSMKGSVRERAEALAHMGSSVFSGITLTK 1206
Qy      1221  LRGILVLGAKAQIQLIPFRMLNLTLGLLGLHGLVPLVPILSVYGPDVNPA 1272
Db      1207  FGGIVVLAKSQIFQIFRYMYLAWVLGATHGLIFLPLVLSYIGPSVKA 1258

```

RESULT 68

```

US-10-995-561-952
; Sequence 952, Application US/10995561
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 952
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-952

```

```

Query Match      34.8%; Score 2402.5; DB 39; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.1e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

```

```

Qy      7  RGWLWALLRLRLAQSEPTTTHQPGYCAFYDCG-----KNPELSSGLMTLSNVCSLS 59
Db      4  RGALGLLALLLCPAQVFSQ-----SCWYGEGLAYGDKRYNCEYISG-----46
Qy      60  NTPARKITGDHILLQKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDN 119

```

```

Db      47  --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKONLQPLQFLSRCPGCFYN 103
Qy      120  FVNHLCHNTCSNQSLFINVTR----VAOLGAGQLPAVVAAYEAFYQHSFAEQSDSCSRV 175
Db      104  LLNLFCELTCSPRQSQFLNVATATEDYVDPVTNCTKTNKVELQYVQGSFANAMYNACRDV 163
Qy      176  RVPAAATLAVGTWCGVYGSGALCNAQRLNFPQGGDTGNGLAPLDIT-----FHLLEPQA 228
Db      164  EAPSSNDKALGLLCKGDADA-CNATNWIYMFKNKGQAPFTITPVPSPDPVH-----215
Qy      229  VSGIQLPNEGVARCNESQGDVATCSODCAASCPAIAIPQ-----ALDSFYLG 279
Db      216  ---GMEPMNATKGCDSVDEVTAPCQDCSIVCGPKQPPPPPPAPWTILGLDAMYVIM 272
Qy      280  QMPGSLVLIILICSVFAVVTILL-----VGFVRVAPARDKSMVDPKGTSL 326
Db      273  WITYMAFLVPGAFPAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy      327  DKLSFSTHTLLGQFFQGGWGTWVASWPLTILVLSVIPVVAALAGLVFTLTDTPVELWSAP 386
Db      325  DPVSAAFEGCLARLFTRMGSFCVRNPGCVIFFSLVFITACSSGLVFRVVTNPDVLSWAP 384
Qy      387  NSQARSEKAPHQHPGPPPTNQVILTAPNRSYRYSLLGPK-NFSGILDLDLLELL 445
Db      385  SSOARLEKEYFDQHPGPFRTQLIIRAPLTDKHIYQYPSPGADVPFGPPLDIQILHQVL 444
Qy      446  ELQERLRLHQLVMSPEAQRNLSLDICYPALNPNTSLYDCCINSLLQYQNNRTLLLLTA 505
Db      445  DLQIAIEN--ITASYDNEVTIQLDCLAPLSPYNT---NCTILSVLNYFQNSHVSIDHKK 499
Qy      506  NOTLMGQTSQVDKDHFLYCANAPLTFKGTALALSCMADYGAPVFPFLAIGYKGDYS 565
Db      500  GDDFF---VYADVHTFLYCVRAPASLNDTSLDHPCLGTFGGVPFVPLVGLGYDDQNYN 556
Qy      566  EBAELIWTPLSNYPAGDPRLAQAKLWEAFLEEMRAPFORRMAGMQVFTFAERSLEDEI 625
Db      557  NATALVITPVVNYNDTEKLQRAQAWKEBFNFVKNYKN---PNLTISFTAERSIEDEL 613
Qy      626  NRTTAEDLPFATSVIYIPLYSALGYSYSSSRVMSVMDSKATLGLGVAVLCAVMAAMG 685
Db      614  NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDKSVLSIGIAGILVLSSVACSLG 673
Qy      686  FFSYLGRISLVILQVVPFLVLSVGADNIPFVLEYQRLPRRPGEPREHIGRALGRVAP 745
Db      674  VFSYIGLPLTLVIEVIFPLVLAAGVDNIFILVQAYORDERLQGETLDQOLGRVLEAVP 733
Qy      746  SMLLCSLSEAI CFFLGALTPMPAVRTPALTSGVLAVILDFLLQMSAFVALLSLDSKQREAS 805
Db      734  SMLFSSFSSTVAFFLGLSMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKRQKKN 793
Qy      806  RLDDVCCVKPQELPPPGQ-GEGLLGFQKAYAPFLLHMTIRGVLLFLFALFGVSLYSM 864
Db      794  RLDFCFCVGAEDGTSVQASESCLFRFFKNSYSPLLLKDMRPVIAIFVGLSFSIAVL 853
Qy      865  CHISVGLDQELALPKDSYLLDYFLNRYFEVGAFFVFTTLLGNFSSBAGMAICSSAG 924
Db      854  NKVDIGLQSLSPDDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKGGQNMVCGMG 913
Qy      925  CNPFSTQKIQVATEPPEQSYLAIPASSWDDPIDWLTP-SSCCRLYISGPNKDKFCPST 983
Db      914  CNNDSLVQOIFNAAQLDNRYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy      984  VNSLNLCKNCSIT-MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db      971  VDDPACVR-CRPLTPEGKORPOGGDFMRFLPMLPFLSDNPNPKCKGGAAYSSAVNTLLGH 1029
Qy      1041  DGQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFEPYPTITNV 1100
Db      1030  GTRVGATYFTVHTVLOTADFDALKKARLIASNT-ETMGINGS--AYRFPFVSFVY 1086
Qy      1101  FFEQYLTLPEGLFMLSCLVPTFAVSCLLLGLDLSGLLNLISIVMILVDTVGFMAIWD 1160
Db      1087  FFEQYLTIIDDTIFNLGVSIGALFVLTWVLGCELWSAVIMCATIAMVLVNMFGVMWLMG 1146

```



```
QY 60 NTPARKITGDHILLOKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPDGVDLVQELCPGFFG-NVSLCCDVRQQLKDNQLPLQFLSRCPSCFYN 103
QY 120 FVNHLCHNTCSNOSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSPEAQSOSCSRV 175
Dd 104 LNLFCELTCSPROQFLNVATEDYVDPVNTQTNVKELOYYVQGSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTCMGVYGVSALCNAQRWLNFOGDTGNGLAPLDIT-----PHLLEPGA 228
Dd 164 BAPSSNDKALGILCGKADDA-CNATNWIYEMFNKONGQAPFTITPVSPDPVH----- 215
QY 229 VGSGLQPLNEGVARCNESQGDVATCSODCAASCPAIARPO-----ALDSTFYLG 279
Dd 216 ---GMEPMNATKGCDESVDVTPCSCDCSIVCGPKPQPPPPAPMTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGFVPAPARDKSKMVDPKGTLS 326
Dd 273 WITYMAFLVFFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFQGWGTWASWPLTILVSLVPPVVALAAGLVFTELTTPDVLWSAP 386
Dd 325 DPVSAAFEGCLRLTRWGSFCVRNPGCVIPFSLVPIFACSSGLVFRVTTNPFVLSAP 384
QY 387 NSQARSEKAFDHOHGFPPRTNQVILTAPNRSYRYDSLLGPK-NFSGILDLLLELL 445
Dd 385 SSOARLEKEYFDHGFPPRTTEQLIIRAPLTDKHIYQYPVPGADVPFGPPLDIQLHVL 444
QY 446 ELQERLRHLQVMSPEAQRNISIQDIQYAPLNPNTSLYDCCINSILQYFQNNRTLLTLTA 505
Dd 445 DLQIAIEN--ITASVDNETVTILQDILAPLSPYNT---NCTILSVLNFQNSHVDHKK 499
QY 506 NOTLMGOTSQVDWKDHFLYCANAPLTFKDGDTALALSCMADYGPVFPPLAIGYKGKDY 565
Dd 500 GDDFF---VYADYHTEFLYCVRAPASLNDTSLHDPCLGTGPGFPVPLVLGYYDDQNTN 556
QY 566 EBAELIMTFLSNLYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVFTTAERSLEDBI 625
Dd 557 NATALVITPPVNNYNDTEKLQRAQAWKEFINVKYKN---PNLITISFTAERSIEDEL 613
QY 626 NRTTADBLDFATSVIVIFLYTSLALGSYSWSRWVDSKATLGGVAVVIGAVMAWG 685
Dd 614 NRESDDVFTVVISYAIMFLYISLALGHKSCRRLLVDSKVSIGTAGILVLSVACSIG 673
QY 686 PFSYLGISSVLQVWPFLVLSGADNIFIFVLEYQRLPRRGPBREVHIGRALGRVAP 745
Dd 674 VFSYIGLPLTLVIEVPELVLAGVDNIFILVQAYQDERLQGETLQOQLGRVITGEVAP 733
QY 746 SMLLCSLSEACIFFLGAITPMPAVRTFALTSLGVLIDPFLQWSAFVALLSDSKRQAS 805
Dd 734 SMFLSSFSETVAFFLGALSVMVAHTFSLFAGLAVFIDPLQITCFVSLILGLDKRQKN 793
QY 806 RLDVCCVKQPLPPPGQ-GEGLILGFQKAVAPFLHWTIRGVVILLFLALFGVSLYSM 864
Dd 794 RLDIFCCVRGADGTSVQASESCLFRFFKNSYSPILLKADWMPRIVIAIPVGLSPSIAVL 853
QY 865 CHISVLQDELALPKDSYLLDYFLFNRYFEVGAIPVYFVTTILGYNFSSEAGMNAICSSAG 924
Dd 854 NKVDIGLQSLMPDSDSYMVDYFKSISQYLHAGPPVYFVLEEGHDYTSKQGNVCGMG 913
QY 925 CNMFSTQKIQVATPEPQESYLAIASSWVDDPIDWLTP-SSCRLXYSGPNKDKFCPS 983
Dd 914 CNNDLSVQOIFNAQQLDNVTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT-MGSVRPSVEQFHYKLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Dd 971 VVDPACVR-CRPLTPEGKQRPQGGDFMFLPMLSDNPNPKCGKGAHAYSSAVNILLGH 1029
QY 1041 DQOVLASRFMATHKPLKNSQDYTEALRAARELANITADLRKVGTDPDAPFVPPYITNV 1100
Dd 1030 GTRVGATFTMYHTVLTQTSADFDALKARLIASNVT-ETMGINGS--AYRVPYSVFV 1086
```

```
QY 1101 FYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDILRSGLNLLSIVMLIVDTVGFMAIWD 1160
Dd 1087 FYEQVLTIDDTIFNLGVSAGAFILVTVWVLGCELWSAVIMCATIAMVLVNMFGVMWLMG 1146
QY 1161 ISYNAVSLINVSAGMSVEFVSHITRSPASTKPTMLERAKEATISMGSAVPAGVAMTN 1220
Dd 1147 ISLNAVSLINVLVMSGIVSEPCSHITRAFTVSMKGSRVEREALAHMSSVFSGITLTK 1206
QY 1221 LFGILVLGLAKAQLIQIFFRNLNLITLGLLHGLVFLFVLVILSYVGPDVNPA 1272
Dd 1207 FGGIVVLAFAKSQIQIFVFRMYLAWVLGATHGLIFLPLVLLSYIGPSVYKA 1258

RESULT 71
US-11-191-274-130
; Sequence 130, Application US/11191274
; GENERAL INFORMATION:
; APPLICANT: Applera Corporation
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LONG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001536PROV
; CURRENT APPLICATION NUMBER: US/11/191,274
; CURRENT FILING DATE: 2005-07-28
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130
; LENGTH: 1278
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-191-274-130

Query Match 34.8%; Score 2402.5; DB 41; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWALLLRLIAQSEPTTTHQPGYCAPYDECG-----KNPELSGLMTLSNVCSLS 59
Dd 4 RGLALGLLGLLCPAQVFSQ-----SCVMYGEGLAYGDKRYNCBSYG----- 46
QY 60 NTPARKITGDHILLOKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDN 119
Dd 47 --PPKPLPDGVDLVQELCPGFFG-NVSLCCDVRQQLKDNQLPLQFLSRCPSCFYN 103
QY 120 FVNHLCHNTCSNOSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSPEAQSOSCSRV 175
Dd 104 LNLFCELTCSPROQFLNVATEDYVDPVNTQTNVKELOYYVQGSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTCMGVYGVSALCNAQRWLNFOGDTGNGLAPLDIT-----PHLLEPGA 228
Dd 164 BAPSSNDKALGILCGKADDA-CNATNWIYEMFNKONGQAPFTITPVSPDPVH----- 215
QY 229 VGSGLQPLNEGVARCNESQGDVATCSODCAASCPAIARPO-----ALDSTFYLG 279
Dd 216 ---GMEPMNATKGCDESVDVTPCSCDCSIVCGPKPQPPPPAPMTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGFVPAPARDKSKMVDPKGTLS 326
Dd 273 WITYMAFLVFFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFQGWGTWASWPLTILVSLVPPVVALAAGLVFTELTTPDVLWSAP 386
Dd 325 DPVSAAFEGCLRLTRWGSFCVRNPGCVIPFSLVPIFACSSGLVFRVTTNPFVLSAP 384
QY 387 NSQARSEKAFDHOHGFPPRTNQVILTAPNRSYRYDSLLGPK-NFSGILDLLLELL 445
Dd 385 SSOARLEKEYFDHGFPPRTTEQLIIRAPLTDKHIYQYPVPGADVPFGPPLDIQLHVL 444
QY 446 ELQERLRHLQVMSPEAQRNISIQDIQYAPLNPNTSLYDCCINSILQYFQNNRTLLTLTA 505
Dd 445 DLQIAIEN--ITASVDNETVTILQDILAPLSPYNT---NCTILSVLNFQNSHVDHKK 499
QY 506 NOTLMGOTSQVDWKDHFLYCANAPLTFKDGDTALALSCMADYGPVFPPLAIGYKGKDY 565
```

```
Db 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLLDHPCLTGCGPVPFVWLVGGYDQNYN 556
Qy 566 EEAALIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFOVTTFAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKQRAQAWKEKFNFVKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSVIVIFLYISLALGSYSSRVMVDSKATLGLGGVAVVLGAVMAAMG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHKSCERLLVDSKVSGLTAGILVLSSVACSIG 673
Qy 686 PFYSLGIRSSLVLQVPPFLVSVGADNIFIPVLEYQRLPRRPBPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVLPFLVAVGVDNIFILVQAYQDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICFFLGALTPMAVTFALTSGLAVIDLFLLOMSAFVALLSLSKROEAS 805
Db 734 SMFLSFSFSTVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIDKROEKN 793
Qy 806 RLDVCCCKVPQELPPPGQ--GEGLLIGFFOKAYAPFLLHWTIRGVVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSVSPLLDKDWMRPVIAIFVGLSFI AVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTILGYNPFSSBAGMNAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEBHDYVTSKGGQNMVCGGGM 913
Qy 925 CNMFSTQKIQYATPEPQOSYLAIAPASSWDDFIDWLT--SSCCRLYISGPNKDECPST 983
Db 914 CNNDLSVQOIFNAQAQLDNYTRIGFAPSSWIDDYFDMVKPQSSCCRY---DNITDQFCNAS 970
Qy 984 VNSLNLCKNMSIT--MGSVRPSVEOFHKYLPWFLDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR--CRPLTEGKQRPQGGDFMFLPFLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGDTPAFEPFYITNV 1100
Db 1030 GTRVGATYEMTYHTVLTQTSADFDALKARLIASNV--ETMGINGS--AYRVFPYSVYV 1086
Qy 1101 FVEOYLTIPEGLFMLSCLVPTFAVSCLLIGLDLSGLNLLSIVMLIVDTVGPMALWD 1160
Db 1087 FVEOYLTIIDDTIFNLGVSGLGAIFLVTMVLGCELWSAIVMCAITAMVLNVNMFVWMLMG 1146
Qy 1161 ISVNAVSLNLVSAVGMSEVESHITRSPAISTKPTWLERAKEATTSMGSAVPAGVAMTN 1220
Db 1147 ISLNAVSLNLVWMSCGISVEFCSHITRAFTVSKGSRVERAEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGLVLGLAKAQLIQIFFRNLNLLITLGLHLGLVFLPVILSVYGVDPVNA 1272
Db 1207 FGSIVVLAFAKSQIFQIFVFRMYLAMVLGLGATHGLIFLPLVLLSYIGPSVYKA 1258
```

RESULT 72

```
US-11-222-021-441
; Sequence 441, Application US/11222021
; GENERAL INFORMATION:
; APPLICANT: Bruno DOMON et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001548
; CURRENT APPLICATION NUMBER: US/11/222,021
; CURRENT FILING DATE: 2005-09-09
; NUMBER OF SEQ ID NOS: 7554
; SOFTWARE: FaetSEQ for Windows Version 4.0
; SEQ ID NO 441
; LENGTH: 1278
; TYPE: PRN
; ORGANISM: Homo sapiens
US-11-222-021-441
```

```
Query Match 34.8%; Score 2402.5; DB 42; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2,3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
Qy 7 RGWLLIALLRLRAQSPYTTIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
```

```
Db 4 RGLALGLLLLLCPAQVFSQ-----SCVMYBEGCIAYGDKRYNCYSG----- 46
Qy 60 NTPARKITGDHLLILLOKICPRLYTPGNTQACCSAKQLVSLBASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG--NVSLCCDVRQLTKDNLQPLQJFLSRCSFCFYN 103
Qy 120 FWNLHCHMTCPNOSLFTNVTR---VAQLGAGQLPAAVAVBAFYQHFQAEQSYDSCSRV 175
Db 104 LLNLFCELTCPRQSQFLNVTATEDYDVPVNTQNTKNKELOYYYVQGSFANAMYNACRDV 163
Qy 176 RVPAAATLAVGTMCVYGSALCNAQRWLNFOGDTGNGLAPLDIT-----PHLLEPGQA 228
Db 164 EAPSSNDKALGLLCKGADA--CNATNWTXYMFKNDNGQAPFTTTPVFSDFPVH----- 215
Qy 229 VGSGLQPLNEGVARCNESQGGDDVATCQDCAASCAPATARPQ-----ALDSTFYLG 279
Db 216 ---GMPEPNNATKGCDESDEVTAFCSCQDCSIVCGPKPQPPPPPPAPMTILGDAMYIM 272
Qy 280 QMPGSLVLIILCSVPAAVVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
Db 273 WITYMAFLVTFGAFVAVWCYRKRYFVSEYTPIDNSIAFSV--NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLLGQFQCGMTWASWPLTILVLVSVIPVVALAAGLVFTTELTTDPDELWSAP 386
Db 325 DPVSAAPBGLRLRLFRMGSCFVRNPGCVIPFSLVFTACSSGLVFVRVTTNPVPLWSAP 384
Qy 387 NSQARSEKAFHDQHPGFFRNTQVLLTAPNRSSRYDLSLLGPK--NFGSILDLDLLELL 445
Db 385 SSQARLEKEYFDQHPGFFRTEQLIRAPLTDKHTYQYPYSGADVPFGPPLDIQLHLQVL 444
Qy 446 ELQERLRHLQVMSPEAQRNISLQDI CYAPLPNDTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVLODILCLAPLSYNT---NCTILSVLNTFQNSHVLDDHK 499
Qy 506 NQTLMGQTSQVMDKDHFLYCANAPLTPKDGITALALSCMADYGA PVPFPLAIGYKGYKS 565
Db 500 GDDPF---VYADYHTHFLYCVRAPASLNDTSLLDHPCLTGCGPVPFVWLVGGYDQNYN 556
Qy 566 EBAALIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFOVTTFAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKQRAQAWKEKFNFVKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSVIVIFLYISLALGSYSSRVMVDSKATLGLGGVAVVLGAVMAAMG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHKSCERLLVDSKVSGLTAGILVLSSVACSIG 673
Qy 686 PFYSLGIRSSLVLQVPPFLVSVGADNIFIPVLEYQRLPRRPBPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVLPFLVAVGVDNIFILVQAYQDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICFFLGALTPMAVTFALTSGLAVIDLFLLOMSAFVALLSLSKROEAS 805
Db 734 SMFLSFSFSTVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIDKROEKN 793
Qy 806 RLDVCCCKVPQELPPPGQ--GEGLLIGFFOKAYAPFLLHWTIRGVVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSVSPLLDKDWMRPVIAIFVGLSFI AVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTILGYNPFSSBAGMNAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEBHDYVTSKGGQNMVCGGGM 913
Qy 925 CNMFSTQKIQYATPEPQOSYLAIAPASSWDDFIDWLT--SSCCRLYISGPNKDECPST 983
Db 914 CNNDLSVQOIFNAQAQLDNYTRIGFAPSSWIDDYFDMVKPQSSCCRY---DNITDQFCNAS 970
Qy 984 VNSLNLCKNMSIT--MGSVRPSVEOFHKYLPWFLDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR--CRPLTEGKQRPQGGDFMFLPFLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGDTPAFEPFYITNV 1100
```

Db 1030 GTRVGATYFMTYHTVLTQTSADFDALKKARLIASNTV-ETMGINGS--AYRVPYSVYV 1086
QY 1101 FYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIWMLVDTVGFMAWMD 1160
Db 1087 FYEQYLTIIIDDTIFNLGVSIGALFLVTMVLGCELWSAVIMCATIAMVLNMFVGMWLMG 1146
QY 1161 ISYNAVSLNLVSAVGMSVEFVSHITRSPAIKPTWLERAKATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKSRVERAEALAHMSSVPSGTLTK 1206
QY 1221 LFGILVLGLAKAQLIQIFPFRMLNLLITLGLHGLVFLPVLVSIVGPDVNP 1272
Db 1207 FGGIVVLAPAKSQIFQIFVFRMYLAWVLGATHGLIFLPLVLSYIGPSVYV 1258
RESULT 73
US-60-423-552-148
; Sequence 148, Application US/60423552
; GENERAL INFORMATION:
; APPLICANT: American Home Product Inc.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING MAJOR DEPRES
; FILE REFERENCE: AM101250L
; CURRENT APPLICATION NUMBER: US/60/423,552
; CURRENT FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 256
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-423-552-148
Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
QY 7 RGWLLWALLRLAQSPPYTHIOPGYCAFVDEG-----KNPELSSGLMTLSNVCSLS 59
Db 4 RGIALGLLGLLLCPAQVFSQ-----SCVMYGECCYAGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLLILQKICPRLTYGNTQACCSAKOLVSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGVLDVQLCPGPFPG-NVSLCCDVRQLTKNLQLOPLQFLSRCPSCFYN 103
QY 120 FVNLHCNTCSNOSLFINVTR----VAOLGAGQLPAVVAYEAFYQHSFAEQSDSCSRV 175
Db 104 LNLFLCELTCSPQSOFLNWTATEDYDVPVNTQKTNVRELQYVYVQSPANMAYNACRDV 163
QY 176 RVPAAATLAVGTCMGVYGSALCNAORWLNFGQDGTGNGLAPLDIT-----FHLLEPGQA 228
Db 164 BAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVPSDPVH----- 215
QY 229 VSGIOPLNEGVARCNESOGDDVATCSQDCAAACPAIARPO-----ALDSTFYLG 279
Db 216 --GMEPMNNAATKGCDEVDVATPCQDCSIVCGPKQPPPPAPWPTILGIDAMYIM 272
QY 280 QMPGSLVLLIILCSFAVVITLL-----VGRVAPARDKSKWDPKGTSL 326
Db 273 WITTYMAFLVFFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSFTHLLGQFGCGWGTWASWPLTILVLSIPVVALAAGLVFTLTTPDVELWSAP 386
Db 325 DPVSAFAFEGCLRLFRMGSPFCVRNPGCVIFPSLVFITACSSGLVFRVTTNPVDLWSAP 384
QY 387 NSQARSEKAFHOGFPGFRTNQVILTAPNRSYRVDLSLLGPK-NFSGILDLDLLELL 445
Db 385 SSQARLEKEYFDFHGFPPFTEQLIIRAPLTDKHIYQPYPSGADVPFGPPLDIQLHQL 444
QY 446 ELQERLRLQVMSPEAQRNLSLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTTLQDICLAPLSPYNT---NCTILSVLNYPQNSHSLDHKK 499

QY 506 NQTLMGQTSQVDKDHFLYCANAPLTFKDGOTALALSCMADYAGAPVPPFLAIGYKGDYS 565
Db 500 GDDFF--VYADYHTHFLYCRAPASLNDTSLLDPLCLGTFGPGVFPFWLVGLGYDQNYN 556
QY 566 EBAELIMTSLNNYPAGDPRLAQAKLEAFLEEMRAFORRMAGMFMQVPTTASRSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLQBAQAEKEFINFVKNYK---PNUTISTASRSLEDEL 613
QY 626 NRTAEDLPFATSYIVIFLYISLALGSSSRVMDSKATGLGAGVAVVLGAVMAAMG 685
Db 614 NRESDSDFVTWISYAIMELYISLALGHSKCRLLVDSKVSIGIAGILVLSVACSLG 673
QY 686 PFSYLGISSVLIVQVVPFLVLSVGADNIFIFVLYEORLPRRPGCPREHIGALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFELVAVGVDNIFILVQAYQORDERLQGETLDOQLGRVLGEVAP 733
QY 746 SMLCSSLSEACIFFICALTPMPAVRTFALTSGIAVLTDLLOMSAFVALLSLDSKQKEAS 805
Db 734 SMFLSFSSTVAFFLGLSVMPAVHTFSLPAGLAVFIDFLQLTCTFVSLGLDIIKQEK 793
QY 806 RLDVCCVKEQBLPPPGQ--GEGILLGFQKAYAPFLHMTTRGVLLLLFLALFGVSLYS 864
Db 794 RLDIFCCVGEAGDTSVQASESCLPRFFKNSYSPLLKDWMRPIVITAFVGVLSPIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYPEVAGPVYFTTGLYFNSSSEAGMNAICSSAG 924
Db 854 NKVDIGLQDLSMPDDSYMVDYFKSISQYLHAGPPVYFVLEEGHDYTSKKGQNMVCGMG 913
QY 925 CNPFSTQKIQYATEPPEQSYLAIPASSHWDDPIDMLTP--SSCCRLYISGPNKDKPCPST 983
Db 914 CNNDLSVQOIFNAAQLDNTRYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQPCNAS 970
QY 984 VNSLNLCKNCSMIT-MGSVRPSVEQPHKYLPLNDRPNIKCPKGLAAAYSTSVN--LTS 1040
Db 971 VVDPAQVR-CRPLTEPGKORPOGGDFMRFLPFLSDNPNPKCGKGGHAYSSAVNILLGH 1029
QY 1041 DQOVLASRPMAYHKLJNSQDYTEALRAARELAANTADLRKVGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADFDALKKARLIASNTV-ETMGINGS--AYRVPYSVYV 1086
QY 1101 FYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIWMLVDTVGFMAWMD 1160
Db 1087 FYEQYLTIIIDDTIFNLGVSIGALFLVTMVLGCELWSAVIMCATIAMVLNMFVGMWLMG 1146
QY 1161 ISYNAVSLNLVSAVGMSVEFVSHITRSPAIKPTWLERAKATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKSRVERAEALAHMSSVPSGTLTK 1206
QY 1221 LFGILVLGLAKAQLIQIFPFRMLNLLITLGLHGLVFLPVLVSIVGPDVNP 1272
Db 1207 FGGIVVLAPAKSQIFQIFVFRMYLAWVLGATHGLIFLPLVLSYIGPSVYV 1258

RESULT 74

US-60-427-579-148
; Sequence 148, Application US/60427579
; GENERAL INFORMATION:

; APPLICANT: American Home Product Inc.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING MAJOR DEPRE
; FILE REFERENCE: AM101250L
; CURRENT APPLICATION NUMBER: US/60/427,579
; CURRENT FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 256
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-427-579-148

Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;

Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECAIYAGDKRYNCBSYG-----46

QY 60 NTPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLEASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRLQTLKDNQLQPLQFLSRCSFCYN 103

QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
DB 104 LNLFCELTCSPRQSOFLNVTATEDYDVPNTQTKNVKELQYVYGQSFANAMYNACRDV 163

QY 176 RVPAAATLAVGTCMGVYGSALCNAQRWLNFGQDTGNGLAPLDT-----PHLEPGQA 228
DB 164 EAPSSNDKALGLLCKGKADA-CNATNWIYEMFNKONGQAPFTITPVFSDFPVH-----215

QY 229 VGSGLQPLNEGVARCNESQGDVATCSCODCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTAPCSCODCIVCGPKQPPPPAPWTLILGLDAMYIM 272

QY 280 QMPSGLVLIILICSVFVVITILL-----VGRVAPARDKSKWVDPKKGTSLS 326
DB 273 WITYMAFLVFPFAGFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

QY 327 DKLSFSTHTLLGQFFQCGGTWASWPLTILVSVIPVVALAAGLVFTELTTDPVELWSAP 386
DB 325 DPVSAAFEGCLRLTRWGSFCVRNPGCVIFFSLVFITACSSGLVFRVVTNTPVDLWSAP 384

QY 387 NSQARSEKAFDHOHGFPPERTNOVILTAENRSYRYSLSLLGPK-NFSGILDLILLELL 445
DB 385 SSQARLEKEYFOHGFPPERTQILRAPLTDKHIYQPYPSGADVPFGPPLDIQILHQLV 444

QY 446 ELQERLRLHQVMSPEARNISLDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPVNT---NCTILSVLNYFQNSHSLDHKK 499

QY 506 NQTLMGQTSQVMDKHFLYCANAPLTFKDGTAALASCMADYAGAPVFPFLAIGYKGYK 565
DB 500 GDDFF---VYADYHTHFLCYVRAPASLNDTSLHDPCLGTFGPGFPVPLVGLGYYDDQYN 556

QY 566 EAEALITPESLNNYPAGDPRLAQALWEAEFALEEMRQFORMMAGMFQVTFRTAERSLEDEI 625
DB 557 NATALVITPVNNYNDTEKLQRAQAEKEFINFKVKNYK---PNLTISFTAERSIEDEL 613

QY 626 NRTTADLPFATSYIVIFLYTSLAGSYSSWSRVWDSKATLGLGVAIVGAVWAAMG 685
DB 614 NRESDDVFTVVISYAIMPLYISLALGHTKSCRRLLVDSKVSIGIAGILIVLSSVACSIG 673

QY 686 PFSYLGRSSLVLQVVPFLVLSVGADNIFIVLEYQRLPRRGPBREVHIGRALGRVAP 745
DB 674 VFSYIGLPLTVLIVIEPFLVLAGVDNIFILVQAYQDERLOGETLQOQLGRVLGEVAP 733

QY 746 SMLLCSLSBAICFFLGALTPMPAVTRFALTGLAVILDLLQMSAFVALLSDSKRQEAS 805
DB 734 SMLSSFSSTVAFGLSVMFAVHTFSLFAGLAVFIDLLAITCFVSLLLGLDIKQEK 793

QY 806 RLDVCCVKQBELPPPGQ-GEGLLLGFFOKAYAPFLLHWTIRGVLLLFLALFGVLSYSM 864
DB 794 RLDIFCVRGAEDGTSVQASESCLFRFFKNYSYSPLLLKDWMRPIVIAIFVGLSPSIAVL 853

QY 865 CHISVLQDELAPKDSYLLDFLPLNRYFVGCAPVYFVTLTGYNFSSEAGNACSSAG 924
DB 854 NKVDIGLDQSLSPDSDSYMDVDFKSIQYLHAGPPVYFVLEEGHDYTSKQGMVCGMG 913

QY 925 CNNFSTQIKQVATEPPEQSYLAIPASSWVDPDFIDWLP-SSCRLYISGPNKDFPCST 983
DB 914 CNNDLSVQIIFNAQDLNTRYTRIGFAPSSWIDDDYFVWVQSSCCRV---DNITDQFCNAS 970

QY 984 VNSLNLCKNMGISIT-MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VVDPACVR-CRLPTPEKGORPQGGDFMRLPMLFSLDNPNPKCGKGAAYSSAVNILLGH 1029

QY 1041 DQVYLASRFMAYHKPLKNSQDYTRALRAARELANITADLRKVPGCTDPAFEVFPYTTITNV 1100
DB 1030 GTRVAGATFMYHTVTLQTSADFDALKKARLIASNVT-ETWINGS--AYRVFPISVFFV 1086

QY 1101 FVEQYLTTLPEGLFMLSCLVPTFAVSCLLLGLDLRLSGLLNLLSIVMLIVDTVGFMAWD 1160
DB 1087 FVEQYLTITIDDTIFNLGVSLSGALFLVTVMLLGCBLWSAVIMCATIAMLVNMGVWLWG 1146

QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFALSTKFTWLERAKEATISMGSAVPAGVAMTN 1220
DB 1147 ISLNAVSLVNLVMSGSISEFCSHITRAFTVSMKGSRYERABEALAHMGSSVFSGITLTK 1206

QY 1221 LPGIILVLGLAKAQLIQIIFPRLNLLITLGLLHGLVFLPVLISVYGGPDVNP 1272
DB 1207 FGGIVVLFAFASQIFQIIFFRMYLAMVLGATHGLIFLPLVLSYIGPSVKA 1258

RESULT 75
US-60-453-050-14124
; Sequence 14124, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14124
; LENGTH: 1278
; TYPE: PR1
; ORGANISM: Homo sapiens
US-60-453-050-14124

Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2,3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECAIYAGDKRYNCBSYG-----46

QY 60 NTPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLEASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRLQTLKDNQLQPLQFLSRCSFCYN 103

QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
DB 104 LNLFCELTCSPRQSOFLNVTATEDYDVPNTQTKNVKELQYVYGQSFANAMYNACRDV 163

QY 176 RVPAAATLAVGTCMGVYGSALCNAQRWLNFGQDTGNGLAPLDT-----PHLEPGQA 228
DB 164 EAPSSNDKALGLLCKGKADA-CNATNWIYEMFNKONGQAPFTITPVFSDFPVH-----215

QY 229 VGSGLQPLNEGVARCNESQGDVATCSCODCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGCDESDEVTAPCSCODCIVCGPKQPPPPAPWTLILGLDAMYIM 272

QY 280 QMPSGLVLIILICSVFVVITILL-----VGRVAPARDKSKWVDPKKGTSLS 326
DB 273 WITYMAFLVFPFAGFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

QY 327 DKLSFSTHTLLGQFFQCGGTWASWPLTILVSVIPVVALAAGLVFTELTTDPVELWSAP 386
DB 325 DPVSAAFEGCLRLTRWGSFCVRNPGCVIFFSLVFITACSSGLVFRVVTNTPVDLWSAP 384

QY 387 NSQARSEKAFDHOHGFPPERTNOVILTAENRSYRYSLSLLGPK-NFSGILDLILLELL 445
DB 385 SSQARLEKEYFOHGFPPERTQILRAPLTDKHIYQPYPSGADVPFGPPLDIQILHQLV 444


```
QY 446 ELQERLRHLQVWSPQAQRNLSLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSYNT---NCTILSVNLVFNQSHSVLDHKK 499
QY 506 NOTLMGQTSQVQWQKDHFLYCANAPLTFKQGTALALSCMADYDGPVFPFLAIGYKGYKDOYS 565
Db 500 GDDFF---VYADYHTHFLYCVAPASLNDTSLLDHDFCLGTFGGVPFWLVGLGYDDQNTN 556
QY 566 EBAALIMTSLNYPAGDPRLAQAALWEBAFLIEMRAFORRMAGMPQVTTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKQRAQAWEKEFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDLPAPATSYIVILYISLALGSYSSRSRWVDSKATLGGVAVVLGAVMAWG 685
Db 614 NRESQSDVTVVVISYAIMFLYISLALGHKSCRRLLVDSKVLGILVILSVVACSILG 673
QY 686 FFSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGPBPREVHIHGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVLAGVDNIFILVQAYQDERLQGETLDQOLGRVLGEVAP 733
QY 746 SMLLCSLSEAICTFFLGALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKREAS 805
Db 734 SMFLSSFSSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIREKQEN 793
QY 806 RLDVCCCVKQBLPPPGQ--GEGLLIGFPQKAYAPFLLHWITRGVULLLFLALFGVSLYSM 864
Db 794 RLDFICCVGAEDGTSTQVASESCLFRFFKNYSPLLLKDMWRPIVIAIFVGLSFIAYL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFVGAUVYFVTTLGYNFSSEAGNACSSAG 924
Db 854 NKVDIGLQDSLSMPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKQGNWCGMG 913
QY 925 CNNFSTQKIQYATEPPEQSYLAIPASSWVDDFDWLTP--SSCCRLYISGPNKDFPCST 983
Db 914 CNNSDLSVQIENAAQLDNYTRIGFAPSSWIDYDFWVKPQSSCCRV---DNITDQPCNAS 970
QY 984 VNSLCLNKNCMGIT--MGSVRPSVEQPHKLPWFLNDRPNIKCPKGLAAYSYVN--LTS 1040
Db 971 VVDPACVR--CRLTPBEGKQRPQGGFMRLPFLMFLSDNPNPKCGKGHAAYSANVILGH 1029
QY 1041 DQOVLASREMAWHPKLNKSQDYTEALRAARELANITADLRKVPGTDPAFVFPYITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADPIDALKKARLASNTV--ETMGINGS--AYRFPYSVYV 1086
QY 1101 FYEQLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSIVMILVDTVGFMAWD 1160
Db 1087 FYEQLTITIDDTIFNLGVSGLAIFLVTVLLGCELWSAIVMCATANVLVNNFGVWLWG 1146
QY 1161 ISYNAVSLNLSVAGMSVEFVSHITRSPAITKPTWLERAKETISMGSAYFAGVAMTN 1220
Db 1147 ISLNAVSLNVLVMSGSGISVEFCSHITRAFTVSMKGSVRERAEALAHMGSSVFSGITLK 1206
QY 1221 LPGAIVLGLAKAQLQIIFFRNLTLITLGLLHGLVFLPVILSYVGPVNPA 1272
Db 1207 FGGIVVLAFAKSIQIFYFRMYLAMVLLGATHGLIFLVLVLSYIGPSVKA 1258

RESULT 76
US-60-453-135-14124
; Sequence 14124, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14124
; LENGTH: 1278
; TYPE: PRN
```

```
; ORGANISM: Homo sapiens
US-60-453-135-14124
```

```
Query Match          34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGLWALLLRLLRLLQAQSEPTTTHQPGYCAFYDEGC-----KNPELSSGLMTLSNVSCLS 59
Db 4 RGLALGLLLLLCPAQFSQ-----SCVMYEGCIAYGDKRYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAQOLVSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFG--NVSLLCCDVRQLQTLKDNLQLPQFLSLRCPSCFYN 103
QY 120 FVNLHCNTCSNQSLFINVTR---VAOLGAGQLPAVVAYBAFYQHSFAEQSYSDSCSRV 175
Db 104 LNLNLCFLTCSPRQSQFLNVTATYDVPVTQTKTNVXELQYYVQGSFANAMYNACRDV 163
QY 176 RYPAATLAVGTMCVGYGSALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLBPGQA 228
Db 164 EAPSNDRKALGLGCKDADA--CNATNWIEMFNKONGQAPFTITPVFSDPPVH-----215
QY 229 VCSGIQPLNEGVARNCSQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVETAPCSCQDCSIVCGPKQPPPPPPAPWTLILGDAMVIM 272
QY 280 OMPGSLVLLIILCSFPAVVITLL-----VGFVAPARDKSKMVDPKKGTSL 326
Db 273 WITYMAFLVFFGAFVAVWCYRKRKFVSEYTPIDSNIAFSV--NASDKG-----ASCC 324
QY 327 DKLSFSTHTLLQFFQGGTGWASWPLTILVLSVIVVVALAAGLVFTELTDPVLEWSAP 386
Db 325 DPVSAAPFGLRLTRWGSFCVRNPGCVIFPSLVPITACSSGLVFRVTTPVDLWSAP 384
QY 387 NSQARSEKAFHDQHPFPFRNTQVILTAENRSYRYSLLGPK--NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDHQFPFRTEQLIRAPLTKHIYQYPYPSGADVFPGLDQILQVQL 444
QY 446 ELQERLRHLQVWSPQAQRNLSLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSYNT---NCTILSVNLVFNQSHSVLDHKK 499
QY 506 NOTLMGQTSQVQWQKDHFLYCANAPLTFKQGTALALSCMADYDGPVFPFLAIGYKGYKDOYS 565
Db 500 GDDFF---VYADYHTHFLYCVAPASLNDTSLLDHDFCLGTFGGVPFWLVGLGYDDQNTN 556
QY 566 EBAALIMTSLNYPAGDPRLAQAALWEBAFLIEMRAFORRMAGMPQVTTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKQRAQAWEKEFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDLPAPATSYIVILYISLALGSYSSRSRWVDSKATLGGVAVVLGAVMAWG 685
Db 614 NRESQSDVTVVVISYAIMFLYISLALGHKSCRRLLVDSKVLGILVILSVVACSILG 673
QY 686 FFSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGPBPREVHIHGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVLAGVDNIFILVQAYQDERLQGETLDQOLGRVLGEVAP 733
QY 746 SMLLCSLSEAICTFFLGALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKREAS 805
Db 734 SMFLSSFSSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIREKQEN 793
QY 806 RLDVCCCVKQBLPPPGQ--GEGLLIGFPQKAYAPFLLHWITRGVULLLFLALFGVSLYSM 864
Db 794 RLDFICCVGAEDGTSTQVASESCLFRFFKNYSPLLLKDMWRPIVIAIFVGLSFIAYL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFVGAUVYFVTTLGYNFSSEAGNACSSAG 924
Db 854 NKVDIGLQDSLSMPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKQGNWCGMG 913
QY 925 CNNFSTQKIQYATEPPEQSYLAIPASSWVDDFDWLTP--SSCCRLYISGPNKDFPCST 983
Db 914 CNNSDLSVQIENAAQLDNYTRIGFAPSSWIDYDFWVKPQSSCCRV---DNITDQPCNAS 970
QY 984 VNSLCLNKNCMGIT--MGSVRPSVEQPHKLPWFLNDRPNIKCPKGLAAYSYVN--LTS 1040
Db 971 VVDPACVR--CRLTPBEGKQRPQGGFMRLPFLMFLSDNPNPKCGKGHAAYSANVILGH 1029
QY 1041 DQOVLASREMAWHPKLNKSQDYTEALRAARELANITADLRKVPGTDPAFVFPYITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADPIDALKKARLASNTV--ETMGINGS--AYRFPYSVYV 1086
QY 1101 FYEQLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSIVMILVDTVGFMAWD 1160
Db 1087 FYEQLTITIDDTIFNLGVSGLAIFLVTVLLGCELWSAIVMCATANVLVNNFGVWLWG 1146
QY 1161 ISYNAVSLNLSVAGMSVEFVSHITRSPAITKPTWLERAKETISMGSAYFAGVAMTN 1220
Db 1147 ISLNAVSLNVLVMSGSGISVEFCSHITRAFTVSMKGSVRERAEALAHMGSSVFSGITLK 1206
QY 1221 LPGAIVLGLAKAQLQIIFFRNLTLITLGLLHGLVFLPVILSYVGPVNPA 1272
Db 1207 FGGIVVLAFAKSIQIFYFRMYLAMVLLGATHGLIFLVLVLSYIGPSVKA 1258
```

```
Db 914 CNNDLSVQOIFNAQAQDNYTRIGPAPSWIDIDYDWMKPSQSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNCNLCNMSIT-MGSVRPSVEQFKYLPWFLNDRNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKORPQGGDFMFLPMFLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADLRKVPQGTDPAFVFPYITNV 1100
Db 1030 GTRVGATYFMYHTVLTQTSADFIDALKARLIASVNT-ETWINGS--AYRVFPYSVFYV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLSTVWLIVDTVGMALWD 1160
Db 1087 FVEQYLTITIDDTIFNLGSLGAIFLVTWVLGCELWSAVIMCATIAMVLVNMFGVWLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERABEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPVILSVYVGPDPNPA 1272
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLAWVLGATHGLIFLPLVLSYIGPSVYKA 1258

RESULT 77
US-60-466-412-14124
; Sequence 14124, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14124
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-14124

Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLRLAQSEPYTHTHQPGYCAFYDCG-----KNPELSGLSMTLSNVCLS 59
Db 4 RGLALGLLILLCLCPAQVFSQ-----SCVMYEGCGIAYGDKRYNCEVSG-----46

Qy 60 NTPARKITGDHILLQKICPRLYTGNTQACCSAKQLVLSLEASISITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDNLQLPLQFLSRCPSCFYN 103

Qy 120 FYNLCHNTCSNQSIFINVTNTR----VAQLGAGQLPAVVAYEAFYQHSFAGSODSCSRV 175
Db 104 LNLNFCELTCSRQSOFLNVTATEDYDVPVNTQTNVKEQLQYVYQGSFANAMYNACRDV 163

Qy 176 RVPAATLAVGTCMGVYGSALCNAORWLNFGQDGTNGLAPLDIT-----PHLLEPGQA 228
Db 164 EAPSSNDKALGLLGGKDA--CNATNWIEMFNKONGQAPFTITPVFSDPPVH-----215

Qy 229 VGSQTOPLNEGVARCNESGGDDVATCSCDCAASCPAIARPO-----ALDSTFYLG 279
Db 216 --GMEPMNATKGCDESDEVTAPCSCDCSIVCGPKPQPPPPAPWTLILGLDAMYIM 272

Qy 280 QMGPSILVLIILCSVPVAVTILL-----VGRVAPARDKSKMVDPKGTSL 326
Db 273 WITYWAFLLVFGAPFAVWCYKRRYFVSEYTPIDNSIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHTLLGQFGQGTWVASWPLTILVLSIPVVAALAGLVFTLTDPVLSWAP 386
Db 325 DPVSAFAEGLRLFRMGSCFVRNPGCVIFLFSVFIATACSSGLVFRVVTINPVDLWSAP 384
```

```
Qy 387 NSQARSEKAFDHFQFPFRTNQVILTAPNRSSYRYSLLLGPK-NPFGILDLDLLLELL 445
Db 385 SSQARLEKEYDQHFQFPFRTNQVILTAPNRSSYRYSLLLGPK-NPFGILDLDLLLELL 444
Qy 446 ELQERLRHLQWMSPEAQRNLSLODICVAPLNPDNTSLYDCCINSLLQYQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSPYNT--NCTILSVNLNFQNSHSLDHKK 499
Qy 506 NQTLMGQTSQVDKDHLYCANAPLTFKDGKTALALSCHADYCAPVPPFLATGGYKGDYS 565
Db 500 GDDEFF--VYADYHTHELYCYRAPASLNDTSLLDHPCLTGFGPPVFWLVGLGYDQDYN 556
Qy 566 EAEALIMTFSLNNYPAGDPRLAQAQAKLEEAFLBEMRAFORRMAGMFOVTFEATERSLEDBI 625
Db 557 NATVALVTFPNNYNTDEKQRAQAKEKEFINFKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLIPATSYIVIFLYISLALGSSYSWSRVMVDSKATLGLGGVAVVLGAYMAAMG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHIKSCRLLVDSKVSGLTAGILVLSVACSLG 673
Qy 686 FFSVLGRTRSSVLIVQVPPFLVLSVADNIFIVLEYQRLPRRPGEPREPHVIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIPFLVAVGVNDNIFLVQAYQRDERLQGETLDQGLGRVLGEVAP 733
Qy 746 SMLJCSLSEATCFELGALTMPAVRTFALTSGLAVILDFLLQMSAFVALLSLDSKRQBAS 805
Db 734 SMFLSSFSSETVAFFLGALUSVMPAVHTLSFLPAGLAVFIDFLQITCFVSELGLDIKQKXN 793
Qy 806 RLDVCCCKVQELPPPGQ-GEGLLLGFFQKAYAPPLLHWITRGVVLVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFPKNSYPLLLKDMWRPIVIAIFVGLVSPSIAVL 853
Qy 865 CHISVGLDQELALPKDSVLLDYFLNRYFVGAPVYVFTVLGYNFSSEAGNAALCSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMDVDFKSIQYLHAGPPVYFVLEEGHDYTSKGNQMVYCGMG 913
Qy 925 CNPFSFTOKIOYATEFPEQSYLAIPASSWVDDFIDWLTP-SSCCRLYISGPNKDFCPSST 983
Db 914 CNDSLVOQIIFNAQAQDNYTRIGPAPSWIDIDYDWMKPSQSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNCNLCNMSIT-MGSVRPSVEQFKYLPWFLNDRNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKORPQGGDFMFLPMFLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADLRKVPQGTDPAFVFPYITNV 1100
Db 1030 GTRVGATYFMYHTVLTQTSADFIDALKARLIASVNT-ETWINGS--AYRVFPYSVFYV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLSTVWLIVDTVGMALWD 1160
Db 1087 FVEQYLTITIDDTIFNLGSLGAIFLVTWVLGCELWSAVIMCATIAMVLVNMFGVWLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERABEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPVILSVYVGPDPNPA 1272
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLAWVLGATHGLIFLPLVLSYIGPSVYKA 1258
```

RESULT 78

US-60-524-882-700

; Sequence 700, Application US/60524882

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: IAKOUBOVA, Olga

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; FILE REFERENCE: CL001494

; CURRENT APPLICATION NUMBER: US/60/524,882

```
; CURRENT FILING DATE: 2003-11-26
; NUMBER OF SEQ ID NOS: 46672
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 700
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-524-882-700

Query Match      34.8%; Score 2402.5; DB 49; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLLAQSEPYTHIQGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DQ 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECAIYAGDKRYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLRSLASITKALLTRCPACSDN 119
DQ 47 --PPKPLPDGVDLVQELCPGFFG--NVSLCCDVRQLQTLKDNLQPLQFLSRCPCSFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAGSDEYSCSRV 175
DQ 104 LNLFLCELTCSPRQSQFLNVTATEDYVDPVTQNTKVELQYVYGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCYVGSALCNAORWLNFGDGTGNGLAPLDIT-----PHLLEPGQA 228
DQ 164 EAPSSNDKALGLLCKDADA-CNATNWIEYMFNKDNGQAPFTITPVFSDFPVH-----215
QY 229 VGSGLQPLNEGVARCNESQGDVATCSQDCAASCPAATAPQ-----ALDSTFYLG 279
DQ 216 ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
DQ 273 WITYWAFLLVPGAFPAWVCYKRYFVSYYTIDSNIAFSV-NAGDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGWGTWASWPLTILVLVSVVVALAAGLVFTLTDPVBLWSAP 386
DQ 325 DPVSAAFEGCLRLFRWGSFCRPNPGCVFVPSLVPITACSSGLVFRVTTNPVLDWSAP 384
QY 387 NSQARSEKAFHQHGGPPRTNQVILTAPNRSRYDLSLLGPK-NFSGILDLLELL 445
DQ 385 SSQARLEKEYFQHFQFPFTEQLIRAPLTDKHIYQYPFGADVPFGPPLDIQLHVL 444
QY 446 ELQERLRHLQVSPRAQRNLSQDICYAPLNPDTSLYDCCINSILQYFQNNRTLLTLTA 505
DQ 445 DLQIAIEN--ITASVDNETVTIQQICLAPLSYNT---NCTLSLVNYPQNSHSLDHKK 499
QY 506 NOTLMGQTSQVDWKDHFLYCANAPLTFKDGITALALS CMADYGA PVFPFLAIGYKGDYS 565
DQ 500 GDDFF--VYADYHTHELYCYVRAPASLNDTSLHDPCLGTGFGVPFVPLVLGSDYDQNVN 556
QY 566 EBAALIMTFLSNYPAGDPRLAQAKLEBAFLBEMRAFQRRMAGMQVFTTFAERSLEDEI 625
DQ 557 NATALVITPPVNNYNDTEKLQRAQAEKEFNFVKYKN----PNLTISFTAESRSTDEEL 613
QY 626 NETTAEDLPIFATSYIVIFLYISLALGSYSSRSRWVDSKATLGLGVAVVLGAWAAG 685
DQ 614 NRESQSDVFTVVISYAIMFLYISLALGHITS CRRLVDSKSLGAGILVILVSSVACSIG 673
QY 686 PFSYLGRISLVILQVPPVLVSVGADNIFIFVLEYQRLPRPFGPREVHIGRALGRVAP 745
DQ 674 VFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAYQDERLOGETDQOLGRVLGEVAP 733
QY 746 SMLLCSLSRAICFFIGALTPMVAFTALTGLAVILDFLQMSAFVALLSDSKRQAS 805
DQ 734 SMFLSSFSFETVAFFLGALSMPAVHTFSLFAGLAFIDFLQITCFVSLGLDIDKQKKN 793
QY 806 RLQVCCCKVQBLPPPGQ--GEGLLIGFFQKAYAPFLLHWITRGVLLFLFALFGVLSYM 864
DQ 794 RLQIFCCVRGABDGTSTVQASESCLFRFPKNSYSPLLKDKMWRPIVIAIFVGLSFI AVL 853
```

RESULT 79

US-60-568-219-476

; Sequence 476, Application US/60568219

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: RESPONSIVENESS TO STATIN TREATMENT, METHODS OF DETECTION AND

; FILE REFERENCE: CL001516

; CURRENT APPLICATION NUMBER: US/60/568,219

; CURRENT FILING DATE: 2004-05-06

; NUMBER OF SEQ ID NOS: 28217

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 476

; LENGTH: 1278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-568-219-476

Query Match 34.8%; Score 2402.5; DB 49; Length 1278;

Best Local Similarity 39.8%; Pred. No. 2.3e-201;

Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLLAQSEPYTHIQGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59

DQ 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECAIYAGDKRYNCEYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLRSLASITKALLTRCPACSDN 119

DQ 47 --PPKPLPDGVDLVQELCPGFFG--NVSLCCDVRQLQTLKDNLQPLQFLSRCPCSFYN 103

QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAGSDEYSCSRV 175

DQ 104 LNLFLCELTCSPRQSQFLNVTATEDYVDPVTQNTKVELQYVYGQSFANAMYNACRDV 163

QY 176 RVPAAATLAVGTMCYVGSALCNAORWLNFGDGTGNGLAPLDIT-----PHLLEPGQA 228

DQ 164 EAPSSNDKALGLLCKDADA-CNATNWIEYMFNKDNGQAPFTITPVFSDFPVH-----215

QY 229 VGSGLQPLNEGVARCNESQGDVATCSQDCAASCPAATAPQ-----ALDSTFYLG 279

DQ 216 ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272

QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSMVDPKKGTSLS 326

```
Db 273 WITYMAFLVFGAFVAVWCYKRYFVSEYTPIDSNIAESV-NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLLGQFFGCGGTWVASWPLTILVSVIPVVALAGLVFTLTTPDVELWSAP 386
Db 325 DPVSAAFEGCLRLRFTRWGSCFVRNPGCVIFFSLVEITACSSGLVFRVTTINPVDLWSAP 384
Qy 387 NSQARSEKAFHDQHGFPFRTNQVILTAPNRSSRYVDSLGLLPK-NFSGILDLLLELL 445
Db 385 SSOARLEKEYFDHGFPGFRTQLIIRAPLTKHIIQYPYPSGADVFPGLDIIQLHQLV 444
Qy 446 ELQERLRHLQVMSPEAQRNISLDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICLAPLSPYNT--NCTILSVLNFQNSHSLVDHKK 499
Qy 506 NQTLMGQTSQVMDKHFLYCANAPLTFKGTALALSCMADYGAPVFPFLAIGGYKGDYS 565
Db 500 GDDFF---VYADYHTHFLYCVRPASLNDTSLLDHPCLGTFGPGVPFVPLVLGGYDQNYN 556
Qy 566 EAEALIMTFSLANNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLOQAQWKEKFINFVNKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTAEDLPFATSYVIFLYISLALGSYSWSRVMVDSKATLGGVAVVLGAVMAANG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHIKSCRLLVDSKVSLSGIAGIILVLSVACSIG 673
Qy 686 FFSYLGIIRSSVLQVVPFLVSYGADNIFIVLEYQRLPRRPGPREVHIQALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIPFLVAVGVDNIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
Qy 746 SMLLCSLSBAICFFLICALTPMPAVRTFALTSGLAVLDFLLQMSAFVALLSDSKROAS 805
Db 734 SMLFSFSTVAFGLGALSVMAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIIRKQKN 793
Qy 806 RLDVCCVKPQBLPPGQ--GEGLLGFFQKAYAPFLHWHITRGVLLFLFALFGVSLYSX 864
Db 794 RLDIFCVRGAEDGTSVQASESCLPRFFKNSYSPLLLLKDMMPVIAIFVGLVSFI 853
Qy 865 CHISVLDELALPKDSYLLDFLFLNRYFVAGAPVYFTTLGYNFSBAGNAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEBHDYTSKQGMVCGMG 913
Qy 925 CNNSFTQIKQVATEPPEOSYLAIAPASSWVDDFIDWLT--SSCCRLYISGPNKDFCPS 983
Db 914 CNNSLVQOIENAAQLDNVTRIGFAPSSWIDDYFDWVKFQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLCLKNVCMGIT-MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR--CRPLTPEGKQPOGGDFMRFLPMFLSDNPNPKCGKGAHAYSSAVNILLGH 1029
Qy 1041 DGOVLASRFMAVHKPLKNSODYTEALRAARELANITADLRKVPDGTDPAFVFPVYITNV 1100
Db 1030 GTRVGATYFMTVHTVLQTSADFDALKKARLASNT--ETWINGS--AIRVFPYSVFV 1086
Qy 1101 FYEQLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSVMLVDTVGFMAWD 1160
Db 1087 FYEQLTIDDIIFNLGSLGAIFLVTWVLLGCELWSAVIMCATAMVLVNMFGVWMLG 1146
Qy 1161 ISYNAVSLNVLVAVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAVAMTN 1220
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVMKGSRVREAEALAHMGSSVFSGITLTK 1206
Qy 1221 LQGLVGLAKAOLQIIFFRNLTLTLGLLHGLVFLPVILSYVGPDPNPA 1272
Db 1207 FGIIVVLAFAKSOIFQIFRYRMLAMVLGLGATHGLIFLPLVLLSYIGPSVKA 1258
```

RESULT 80

US-60-592-195-129

; Sequence 129, Application US/60592195

; GENERAL INFORMATION:

; APPLICANT: Applera Corporation

```
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01536PROV
; CURRENT APPLICATION NUMBER: US/60/592.195
; CURRENT FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 1278
; TYPE: PRN
; ORGANISM: Homo sapiens
; US-60-592-195-129
```

```
Query Match 34.8%; Score 2402.5; DB 49; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
Qy 7 RGWLLWALLRLLAQSEPYTTIHQPCYCAFYDBCG-----KNPELSGLMTLSNVSCLS 59
Db 4 RGLALGULLLLLCPAQVFSQ-----SCVWYCEGCIAYGDKRYNCEYSG-----46
Qy 60 NTPARKITGDHLIILQKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDN 119
Db 47 --PKPLPKDGYDLVQELCPGFFPG-NVSLCCDVRLQTLKDNLQLPLQFLSRCPSCFYN 103
Qy 120 FVNLHCHNTCSPNOSLFINVTR----VAQLGAGQLPAVVAEAFYQHSFASQSYDCSRV 175
Db 104 LNLNFCULTCSPROSQFLNVTATEDYVDPTNQTNTNKEIQYVVGQSFANAMYNACRDV 163
Qy 176 RVPAAATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDIT-----PHLLEPGA 228
Db 164 EAPSSNDKALGLCGKADA--CNATNWLEYMFKNDNGOAPFTIIPVFSDFPVH-----215
Qy 229 VSGIQLPNEGVARCNESQGDVATCSCQDCAACPAIARPO-----ALDSTFVLG 279
Db 216 ---GMPEMNNATKGDSEVDVETAPCSCQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
Qy 280 QMPGSLVLIILCSVFAVVITLL-----VGRVAPARDKSMVDPKKGTSLS 326
Db 273 WITYMAFLVFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLLGQFFGCGGTWVASWPLTILVSVIPVVALAGLVFTLTTPDVELWSAP 386
Db 325 DPVSAAFEGCLRLRFTRWGSCFVRNPGCVIFFSLVEITACSSGLVFRVTTINPVDLWSAP 384
Qy 387 NSQARSEKAFHDQHGFPFRTNQVILTAPNRSSRYVDSLGLLPK-NFSGILDLLLELL 445
Db 385 SSOARLEKEYFDHGFPGFRTQLIIRAPLTKHIIQYPYPSGADVFPGLDIIQLHQLV 444
Qy 446 ELQERLRHLQVMSPEAQRNISLDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICLAPLSPYNT--NCTILSVLNFQNSHSLVDHKK 499
Qy 506 NQTLMGQTSQVMDKHFLYCANAPLTFKGTALALSCMADYGAPVFPFLAIGGYKGDYS 565
Db 500 GDDFF---VYADYHTHFLYCVRPASLNDTSLLDHPCLGTFGPGVPFVPLVLGGYDQNYN 556
Qy 566 EAEALIMTFSLANNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLOQAQWKEKFINFVNKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTAEDLPFATSYVIFLYISLALGSYSWSRVMVDSKATLGGVAVVLGAVMAANG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHIKSCRLLVDSKVSLSGIAGIILVLSVACSIG 673
Qy 686 FFSYLGIIRSSVLQVVPFLVSYGADNIFIVLEYQRLPRRPGPREVHIQALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIPFLVAVGVDNIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
Qy 746 SMLLCSLSBAICFFLICALTPMPAVRTFALTSGLAVLDFLLQMSAFVALLSDSKROAS 805
Db 734 SMLFSFSTVAFGLGALSVMAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIIRKQKN 793
```



```

Db      216  ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKQPPPPPPAPWTLILGLDAMYVM 272
QY      280  QMPSGLVLIILCISVFAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
Db      273  WITYMAFLVFFGAFVAVCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY      327  DKLSSTHTLLQFQCGWGTWASWPLTLVLVSVVVALAAGLVFTELTTDPVELWSAP 386
Db      325  DVSAAFEGLRLRLETRMGSFCVRNPGCVIFPSLVPITACSGLVFVRVTTPVDLWSAP 384
QY      387  NSQARSEKAFDQHFQFPFRTNQVILTAPNRSSRYDLSLLGPK-NFSGILDLILLELL 445
Db      385  SSQARLEKEYFDQHFQFPFRTEQLIRAPLTDKHIYQYPSPGADVPFGPPLDIQLHQL 444
QY      446  ELQERLRHLQVMSPEAQRNLSQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLTA 505
Db      445  DLQIAIEN--ITASVDNETVTLDICLAPLSPYNT---NCTILSVLNFQNSHVLHKK 499
QY      506  NOTLMGQTSQVDKDHFLYCANAPLTFKDGTAALASCMADYCAPVPPFLAIGYKGYDS 565
Db      500  GDDFF---VYADYHTHFLYCVRAPASLNDTSLLDHPCLGTFGCPVPMVLVGLGYDQNYN 556
QY      566  BAEALIMTFSLNYPAGDPRLAQAKLWEAEFLERAFQRRMAGMFQVTFTAERSLEDEI 625
Db      557  NATALVITFPVNNYNDTEKQRAQAKEFNFVKNYKN---PNLTISFTAERSIEDEL 613
QY      626  NRTTAEDLPFATSYIVIFLYISLALGSYSMSRVWVDSKATLGGVAVVLGAVMAANG 685
Db      614  NRESDSVFTVVISYAIMFLYISLALGHIKSCERLLVDSKVSIGIAGILVLSSVACSLG 673
QY      686  FFSYLGRSSVLVQVFPFLVLSVADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db      674  VFSYIGLPLTLVIEVPELVLAVGVDNIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
QY      746  SMLLCSLSEAICFFLIGALTTPMAVRTFALTSLGLHVLDPVLQMSAFVALLSLDSKQRAS 805
Db      806  RLDDVCCVKPQLPPPGQ-GEGLLGFQKAVAPFLHWHITRGVLLFLFALFGVLSYM 864
Db      794  RLDFICCVRGADGTSVQASESCLFRFFANUSPLLLKDWMPRIVIALFVGLSFI 853
QY      865  CHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYFVTLLGYNFSSEAGMNAICSSAG 924
Db      854  NKVDIGLQSLMPDSDSYWDYFKSISQYLHAGPPVYFVLEGRHDYTSKQGNVYCGMG 913
QY      925  CNPFSFTQKIQVATEFPQSYLAIPASWVDDFDWLTP-SSCRLYISGPNKDKPCPST 983
Db      914  CNNDSLVQOIFNAQAOLDNVTIRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY      984  VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db      971  VDPACVR-CRLTPEGKORPGGDFMFLPMLSDNPNKCGKGHAAYSAVNILLGH 1029
QY      1041  DQCVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITNV 1100
Db      1030  GTRUGATYFTWYHTVLTQTSADRIDALKARLIASNT-ETWINGS--AYRVFPYSVFV 1086
QY      1101  FYEQLYTLPEGLFMLSCLVPTPAVSCLLGLDRSLGNLLLSIVMLVDTVGPMAWMD 1160
Db      1087  FYEQLYTLTIDDTIFNLGVSIGAIFLVMTVWLLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
QY      1161  ISYNAVSLINLVSAGMSVEFVSHITRFSATSKPTWLERAKEATISMGSAVFAVAMTN 1220
Db      1147  ISLNAVSLVNLVMSGIGSVEFCSHITRAFTYPMKGSRYERAEALAHGSSVFSGITLTK 1206
QY      1221  LPGAIVLGLAKAQLQIIFERLNLTLITLGLLHGLAVFLPVILSYVGPDPVWPA 1272
Db      1207  FGGIVVLFAFASQIFQIFFRMYLANVLLGATHGLIFLPLVLSYIGSPSNKA 1258

```

RESULT 84

```

US-60-680-002-1038
; Sequence 1038, Application US/60680002
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001601
; CURRENT APPLICATION NUMBER: US/60/680,002
; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1038
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-1038

Query Match      34.8%; Score 2402.5; DB 50; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY      7  RGWLLWALLRLIAQSEPYTTTHQPGYCAYFDECG-----KNPELSGSLMTLSNVSCLS 59
Db      4  RGLALGLLILLLLCPAQVFSQ-----SCVMYGEGLAYGDKRYNCEYSG-----46
QY      60  NTPARKITGDHILILLOKICPRLYTGPNTQACCSAKQLVSLEASLSITKALLTRCPACSDN 119
Db      47  --PPKPLPKDGYDLVQELPCPGFFFG-NVSLCCDVRQLQTLKDNQLQLPLQFLSRCPSCFYN 103
QY      120  FVNHLCHNTCSPNOSLFINVTR-----VAQLNGAQPLPAVVAEAFYQHSFAEOSYDSCSRV 175
Db      104  LLNLFCELTCSPROQFLNVTATEDYVDPVTNQTNTKVELQYVYQGSFANAMYNACRDV 163
QY      176  RVPAATLAVGTMCCVYGSALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLEPGQA 228
Db      164  EAPSSNDKALGLLCKSDADA-CNATNWIYFWPKNGQAPPTITVPFSDPPVH-----215
QY      229  VSGSQPLNEGVARNCSQGDVATPCSCDCAASCAIARPO-----ALDSTFYLG 279
Db      216  --GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKQPPPPPPAPWTLILGLDAMYVM 272
QY      280  QMPSGLVLIILCISVFAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
Db      273  WITYMAFLVFFGAFVAVCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY      327  DKLSSTHTLLQFQCGWGTWASWPLTLVLVSVVVALAAGLVFTELTTDPVELWSAP 386
Db      325  DVSAAFEGLRLRLETRMGSFCVRNPGCVIFPSLVPITACSGLVFVRVTTPVDLWSAP 384
QY      387  NSQARSEKAFDQHFQFPFRTNQVILTAPNRSSRYDLSLLGPK-NFSGILDLILLELL 445
Db      385  SSQARLEKEYFDQHFQFPFRTEQLIRAPLTDKHIYQYPSPGADVPFGPPLDIQLHQL 444
QY      446  ELQERLRHLQVMSPEAQRNLSQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLTA 505
Db      445  DLQIAIEN--ITASVDNETVTLDICLAPLSPYNT---NCTILSVLNFQNSHVLHKK 499
QY      506  NOTLMGQTSQVDKDHFLYCANAPLTFKDGTAALASCMADYCAPVPPFLAIGYKGYDS 565
Db      500  GDDFF---VYADYHTHFLYCVRAPASLNDTSLLDHPCLGTFGCPVPMVLVGLGYDQNYN 556
QY      566  BAEALIMTFSLNYPAGDPRLAQAKLWEAEFLERAFQRRMAGMFQVTFTAERSLEDEI 625
Db      557  NATALVITFPVNNYNDTEKQRAQAKEFNFVKNYKN---PNLTISFTAERSIEDEL 613
QY      626  NRTTAEDLPFATSYIVIFLYISLALGSYSMSRVWVDSKATLGGVAVVLGAVMAANG 685
Db      614  NRESDSVFTVVISYAIMFLYISLALGHIKSCERLLVDSKVSIGIAGILVLSSVACSLG 673
QY      686  FFSYLGRSSVLVQVFPFLVLSVADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db      674  VFSYIGLPLTLVIEVPELVLAVGVDNIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
QY      746  SMLLCSLSEAICFFLIGALTTPMAVRTFALTSLGLHVLDPVLQMSAFVALLSLDSKQRAS 805

```



```
Db 734 SMFLSSFTVAFFLGLSVMPAVHTSLFAGLAVFIDFLQITCFVSLGLDIDKQKCN 793
Qy 806 RLDVCCVKPQELPPPGQ-GBGLLGGFQKAYAPFLHLHWITRGVLLFLFALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLKDKWMPVIAIFVGVLSFSAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPVYFVTLGYNFSSAGHNAICSSAG 924
Db 854 NKVDIGLQSLMPPDDSYMVDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVCQGMG 913
Qy 925 CNMFSTQKIQVATEFPFQSYLAIPASSWDDFIDWLT-SSCCRLYISGPNKDFPCPT 983
Db 914 CNNSLVQOIFNAALQDNYTRIGFAPSSWIDDYFDMVAPQSSCCRV---DNITQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFLNDRNIPKCGGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSASVNLIGH 1029
Qy 1041 DGOVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADPIDALKKARLIASNT-ETWINGS--AYRVFPYSVFYV 1086
Qy 1101 FVEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWLD 1160
Db 1087 FVEQVLTITDDITFNLGVSLSGAIFLVWVLLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
Qy 1161 ISYNAVSLINLSVAGMVSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGIVSEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPVILSYVGPVNPA 1272
Db 1207 FGGIVVLAPAKSQIFQIFRMYLAMVLLGATHGLIFLVLVSYIGPSVNKA 1258

RESULT 85
US-60-720-274-191
; Sequence 191, Application US/60720274
; GENERAL INFORMATION:
; APPLICANT: Olga Iakoubova
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Cardiovascular Disorders and Drug Response, Methods of
; TITLE OF INVENTION: Detection and Uses Thereof
; FILE REFERENCE: CD000002
; CURRENT APPLICATION NUMBER: US/60/720,274
; CURRENT FILING DATE: 2005-09-23
; NUMBER OF SEQ ID NOS: 22965
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 191
; LENGTH: 1278
; TYPE: PRN
; ORGANISM: Homo sapiens
US-60-720-274-191

Query Match 34.8%; Score 2402.5; DB 51; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 40; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLAQSEPYTHIQPGYCAFYDCG-----KNPELSGSLMTLSNVSCLS 59
Db 4 RGLAUGLLLLLCPAQFSQ-----SCWYGBGCIAYGDKRYNCEVSG-----46
Qy 60 NTPARKITGDHILLOKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRLQTLKDNLQLPLQLSRCPSCFN 103
Qy 120 FVNLHCNTCSNQSLFINVTR----VAOLGAGQLPAAVAYEAFYQHSFAEYSQSCSRV 175
Db 104 LLNLFCELTCSQRSQFLNVTATEDYVDPVNTQTKNVKELQYVVGQSPANAMYNACRDV 163
Qy 176 RVPAAATLAVGTCVYGYSALCNAQRWLNPFQDGTGNGLAPLDIT-----PHLEPGQA 228
```

```
Db 164 EAPSSNDKALGLLCKDADA-CNATNMEYFNKONGQAPFTITVPFSDFPVH-----215
Qy 229 VSGIQPLNEGVARCNESQGDVATCSQDCQCAASCPAIARPO-----ALDSTFYLG 279
Db 216 ---GMBEMNNATKGBDESVDDEVATPACSDQCSICVCGPKPQPPPPAPMTIILGLDAMYIM 272
Qy 280 QMPCSLVLIILCSVEAVVTILL-----VGFRVAPARDKSMVDPKKGTSL 326
Db 273 WITMAFLVFPFGAFFAVWCVRKRYFYSEYTPIDSNATFSV-NASDKGE-----ASCC 324
Qy 327 DKLSEFSTHLLGQPFQKGWTVASWPLTILVLSVIPVVALAAGLVFTELTITDTPVELWSAP 386
Db 325 DPVSAAEFEGCLRLFTRWGSCFVNPVGVIFSVLFTACSSGLVFVRVTINPVDLWSAP 384
Qy 387 NSQARSKAFHQHGFQPFRTNQVILTAPNRSSRYDSLLLGPK-NFSGILDLDLLELL 445
Db 385 SSQARLEKEYDQHFQFPFRTEQLIIRAPLTDKHIYQYPYSGADVPFGPPLDIQLHQVL 444
Qy 446 ELQERLRLHQLVMSPEAQRNISLQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLTA 505
Db 445 DLQIAIEN--ITASDYNETVILQICLAPLSFYNT---NCTILSVLNTFQNSHSLVDHKK 499
Qy 506 NOTLMGOTSQVWDKHDFLYCANAPLTFKDGFTALALSCMADYGA VFPFLAIGGYKQDYS 565
Db 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLLDHDPCLGTGPGVPFVWLVGGYDQNYN 556
Qy 566 EBAELIMTFSLNNYPAGDPRLAQAKLWEEAELEEMRAFQRRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPPNNYNTDETEKLQRAQMEKBFNFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPIFATSVIVILYISALGYSWSRVMVDSKATLGLGGVAVVLGAVMAAG 685
Db 614 NRESDDVFTVVISALMFLYISIALGHIKCRULLVDSKVSGLAGILIVLSSVACSLG 673
Qy 686 PFSYLGIRSLVILQVVPFLVSLVSGADNIFVLEYQRLPRRPPREPHVIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFVLAVGVNDIFILVQAYQDERLQGTTLQQLGRVLGEVAP 733
Qy 746 SMLCSSLSEATCFPLGALTPNAVTRFALTSGLAVIDFLQMSAFVALLSIDSROEAS 805
Db 734 SMFLSSFTVAFFLGLSVMPAVHTSLFAGLAVFIDFLQITCFVSLGLDIDKQKCN 793
Qy 806 RLDVCCVKPQELPPPGQ-GBGLLGGFQKAYAPFLHLHWITRGVLLFLFALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLKDKWMPVIAIFVGVLSFSAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPVYFVTLGYNFSSAGHNAICSSAG 924
Db 854 NKVDIGLQSLMPPDDSYMVDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVCQGMG 913
Qy 925 CNMFSTQKIQVATEFPFQSYLAIPASSWDDFIDWLT-SSCCRLYISGPNKDFPCPT 983
Db 914 CNNSLVQOIFNAALQDNYTRIGFAPSSWIDDYFDMVAPQSSCCRV---DNITQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFLNDRNIPKCGGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSASVNLIGH 1029
Qy 1041 DGOVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADPIDALKKARLIASNT-ETWINGS--AYRVFPYSVFYV 1086
Qy 1101 FVEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWLD 1160
Db 1087 FVEQVLTITDDITFNLGVSLSGAIFLVWVLLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
Qy 1161 ISYNAVSLINLSVAGMVSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGIVSEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFLPVILSYVGPVNPA 1272
```



```
Db 1207 FCGIVLAPAKSQIFQIFYFRMYLWMLVLLGATHGLIFLPVLLSYIGSVNKA 1258
RESULT 86
US-11-222-021-440
; Sequence 440, Application US/11222021
; GENERAL INFORMATION:
; APPLICANT: Bruno DOMON et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001548
; CURRENT APPLICATION NUMBER: US/11/222,021
; CURRENT FILING DATE: 2005-09-09
; NUMBER OF SEQ ID NOS: 7554
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 440
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-222-021-440
Query Match 34.8%; Score 2402.5; DB 42; Length 1279;
Best Local Similarity 39.8%; Pred. No. 2.4e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
QY 7 RGMWLLALLRLRLAQSBPYTHIQGYCAFVDECG-----KNPELSSGSLMTLSNVCSLS 59
Db 4 RGLALGALLLLLLCPAQVFSQ-----SCWYGEIGYAGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQIVLSLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKONLQPLQFLSRCPSCFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR----VAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCELTCSPRSQSLFNVATATEDYDPTNQTNTKTNKVELQYVYVGSFANMYNACRDV 163
QY 176 RVPAATLAVGTWCGYVGSALCNAORLWLNFGSDTGNGLAPLDIT-----PHLLBPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIETMFKNDGQAPFTITPVFSDFPVH----- 215
QY 229 VSGIOPLEAGVARNCHESQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 --GMEPMNATKGCDESDEVTPACSCQDCSIVCGKFPQPPPPAPWTILGLDAMVYIM 272
QY 280 QMPGSLVLIILCSVFAVNTILL-----VGRVAPARDKSMVDKPKGTSL 326
Db 273 WITYMAFLVFFGAFPAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGQGWGTWASWELTILVLSVIPVVALAAGLVFTLTDPVELWSAP 386
Db 325 DPVSAAFEGCLRLFRWGSFCVRNFGCVIFFSLVFTACSSGLVFRVVTNPDVLSAP 384
QY 387 NSQARSEKAFDHQFGPFRFTNQVIITAPNRSRYVDSLLGPK-NFSGILDLDLLELL 445
Db 385 SSQARLEKEYFDQHFQGFRTTEQLIIRAPLTDKHIQYPYPSGADVPFGPDLQIILQVL 444
QY 446 ELQERLRHLQVMSPEARNLSDICVAPLNDNTSLYDCCINSLLQYQNNETLILTTA 505
Db 445 DIQIAIEN--ITASYNDETITQDICIAPLSPYNT---NCTLSVLNLYFQNSHVLDDHK 499
QY 506 NOTLMQTSQVQWQKDFLYCANAPLTFKDGITALALSCMADYAGAPVPFPFLAIGYKGOYS 565
Db 500 GDDFF---VYADVHTFLYCVRAPASINDTSLHDFCLGTFGGVPFWLVGLGYDDQNTN 556
QY 566 EBAELMTFSLNYPAGDPRLAQKLWEBAFLEEMPAFORMMAGMPQVTTFAERSLEDEI 625
Db 557 NATALVITPVNYYNDTEKLQRAQAWKEKFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEPLPIFATSYIVIFLYTSLALGYSYSSSRVMDSKATILGCGVAVVLGAVMAWG 685
Db 614 NRESDSVFTVWVSYAIMFLYISLALGHITKSCRRLLVDKSVISGIAGILIVLSSVACSLG 673
QY 686 FFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHVHIGRALGRVAP 745
Db 674 VESYIGLPLTLVIEVIFPLVLAAGVDNIFILVQAYQDRERLQGETILDQQLGRVLGVAP 733
QY 746 SMLLCSLSRAICFFLGCALTPMPAVRTPALTSGLAVIDLPLOMSAFVALLSLSKQREAS 805
Db 734 SMLSSFSFSETVAFPLGALSVMPPAVHTFSLFAGLAVFIDFLQLITCFVSLILGLDKQKEKN 793
QY 806 RLDVCCCVKPELPPPGQ-GEGLLGFQKAYAPFLLHMTITRGVWLLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFKNSYSPLLKKOMMRPIVIAIFGVLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYPFLNRYFEVGAPEVYVTTLYGNFSSEAGMAICSSAG 924
Db 854 NKVDIGLQDLSMPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDYTSSTSKGQNMVCGMG 913
QY 925 CNNFSTQKIQVATEPPEQSILAIAPASSWDDDFIDMLTP--SSCCRLYISGPNKDKCPCST 983
Db 914 CNNDLSVQQLFNAQLDNTRYTRIGFAPSSWIDDFDMVKPOSSCCRV---DNITDQPCNAS 970
QY 984 VNSLNLCKNCMSIT-MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLLAAYSTSVN--LTS 1040
Db 971 VVDPAVCR-CRPLTPGKQRPQGGDFWRFLPMLPSNDPNPKCGKGHAAVSSAVNILLGH 1029
QY 1041 DGOVLASRRPMAYHKPLKNSQDYTEALRAARELANITADLRKVPETDPAFEVPPYITNV 1100
Db 1030 GTRVGATYPMYVHTVLTQTSADFDALKARLIASNVT-ETMGINGS--AVRVFPYSVFYV 1086
QY 1101 FRYEYLTIIPEGLFMLSLCLVPTFPVAVSCILLGLDLSGLNLLLSIVMILVDTVGFMAIWD 1160
Db 1087 FRYEYLTIIIDDYIFNLGSLGAIFLVTWLLGELWSAIVMCATIAMVLNMGVMMWLMG 1146
QY 1161 ISYNAVSLNLVSAVGMSEVFSVSHITRSPAISTKPTWLERAKEATTSMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLWMSGIVSEVFCSHITRAFTVSKGSRVERAEALAHMGSSVSGITLTK 1206
QY 1221 LRGILVLGLAKAQLIQIFPFRNLMLTLTLGLHLVFLPVLVSYVSGPDVNP 1272
Db 1207 FCGIVVLAFAKSGIFQIFYFRMYLWMLVLLGATHGLIFLPVLLSYIGSVNKA 1258
RESULT 87
US-60-680-002-1037
; Sequence 1037, Application US/60680002
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001601
; CURRENT APPLICATION NUMBER: US/60/680,002
; CURRENT FILING DATE: 2005-05-12
; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1037
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-1037
Query Match 34.8%; Score 2402.5; DB 50; Length 1279;
Best Local Similarity 39.8%; Pred. No. 2.4e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
QY 7 RGMWLLALLRLRLAQSBPYTHIQGYCAFVDECG-----KNPELSSGSLMTLSNVCSLS 59
Db 4 RGLALGALLLLLLCPAQVFSQ-----SCWYGEIGYAGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQIVLSLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKONLQPLQFLSRCPSCFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR----VAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCELTCSPRSQSLFNVATATEDYDPTNQTNTKTNKVELQYVYVGSFANMYNACRDV 163
```


Db 602 ALVITFPVNNYNDTERLQRAWKEFISFVKYKN---PNUITISFTEARSIEDENRE 658
 QY 629 TAEDLPATSYIVIFLYISLALGSYSSMRVMDSKATLGLGVAVVLGVANMAAGPFS 698
 Db 659 SNSDVFVTVIISVYVMEFLYISLALGHISQSRLLVDKISLGIAGILVSSVACSLGIFS 718
 QY 689 YLGIRSLVLOVVPFLVSVGADNIFIFVLEYQRLPRRPGEPREHIGRALGRVAPSM 748
 Db 719 YMGEMPTLIVIEVIFPLVAVGVDNIFILVQYQERLQBEETLDOOLGRILEVAPTWF 778
 QY 749 LCSLSAICFFLGAITPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKROEASRLD 808
 Db 779 LSPFSETSAFFGALGSMFAVHTFSLFAGMAVLIDFLQLTCTVSLGLDIDKQKHL 838
 QY 809 VCCVPRQELPPPGQ-----EGLLGFQKAYAPFLHMITRQVLLLFLALFGVLSYM 864
 Db 839 ILCVRGAD--DGQSHASESYLFRFKNYPAPLKKLKWLRPIVAVFVGLVSPSAVV 895
 QY 865 CHISVGLDQELALPKDSYLLDYFLNRPFEVGAFFVFTTLGYNFSSEAGMAICSSAG 924
 Db 896 NKVDIGLDQSLMPNDSYVIANFKSLAQYLHSGPPVYFVLEBGNYSRRKGQNMVCGMG 955
 QY 925 CNNFSFTQIOYATEPEPOSYLAIPASSWVDDFDIMLTP--SSCCRLYISGPNKDFCRPT 983
 Db 956 CNDLSVQQLFNAEALDITTRVGFAFSSWIDDFWVSQSSCCRLY---NVTHQFCNAS 1012
 QY 984 VNSLNLKNCMSIT--MGSVRPSVEQFKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDG 1042
 Db 1013 VMDPTCVR--CRPLTPGKQRPQKPFMKFLPMLSDNPNPKCGKGGHAYGSAVINVGDD 1071
 QY 1043 QVL--ASRFMAYKPLKNSQDYTBALRAARELANITADLRKVGTDPAPFVPPYITNVF 1101
 Db 1072 TVIGATYFTYTHILKTSADYTDAMKARLIASNITETWRS--KGSD--YRVFPYSVYVF 1128
 QY 1102 YEQYLTLPEGLPMLSLVPTFAVSCLLGLDLRSLNLISLVMLIVDTVGFMAIWDI 1161
 Db 1129 YEQYLTIIDTIFNLVSLSGSIPLVTLVGLCELWSAVIMCTIAMILVNMFGVMWLGI 1188
 QY 1162 SYNNAVSLNLVAVGMSVFSVSHITRSPFAISTKPTWLERAKEATISMGSAVAGAVMTNL 1221
 Db 1189 SLNAVSLNLVMSGSI SVFCSHITRAFTWSTKGSRSVRAERLAHMGSVSGITLTKF 1248
 QY 1222 PGLVGLAKAQIQLIFFRLNLLITLGLLHGLVFLPVLVSYVGPDPNPA 1272
 Db 1249 GGVIVLAFAKSQIFEIFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVKA 1299

RESULT 89
 PCT-US03-24982A-291
 ; Sequence 291, Application PC/TUS0324982A
 ; GENERAL INFORMATION:
 ; APPLICANT: Syngenta Participations AG
 ; APPLICANT: Stam, Lynn
 ; APPLICANT: Kamdar, Kim
 ; APPLICANT: Spana, Eric
 ; APPLICANT: Bachmann, Jane
 ; TITLE OF INVENTION: Nucleic Acid Sequences from Drosophila Melanogaster that Encode H
 ; FILE REFERENCE: 70131WOPCT
 ; CURRENT APPLICATION NUMBER: PCT/US03/24982A
 ; CURRENT FILING DATE: 2003-08-08
 ; PRIOR FILING DATE: 2002-10-29
 ; NUMBER OF SEQ ID NOS: 381
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 291
 ; LENGTH: 1287
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 PCT-US03-24982A-291

Query Match 32.7%; Score 2262; DB 1; Length 1287;
 Best Local Similarity 37.8%; Pred. No. 6.1e-189;

Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;
 QY 10 LLMALILRLAQSEPYTTIHQPGYCAFYDCGKPNELSGSLMTLSNVSCLSNTPARKITGD 69
 Db 17 LIAAULPTLIQSSKOD-----CWYGVNTN-DESHS-----QNCYPNGTAKEMAYD 62
 QY 70 HLILQKICPLRYTGNTQACSAKQLVSLASLSITKALLTRCPACSNFVNLHCHNTC 129
 Db 63 GLELLKRCGFLLESENKFCDDKNQVELLNKVELAGNILDRCPCSMENLVRIHQFTC 122
 QY 130 SPNOSLFINVTRVAQLGAGLPAVVAEAPYQHSFAEQSYDSCSRVRVPAATLAVGTWC 189
 Db 123 SPQRAEFMHVATQKNKKGD--EYISSVDLHISTEYINKTKYSQSVSPQTQGLAFDLMC 181
 QY 190 GYVGSALCNAQRLWLFQGTGGLAPLDITFHLRPGQAVGSGIQIQLNEGVARNESQGD 249
 Db 182 GAYSASRCNPTKWFNMGDATNYPVPQIYIQHEP--KSNNNFTPLANTVTPCQAVSS 240
 QY 250 DVATSCQDCAASC-----PAIARPO-----ALDSTFYLQMPGSLVLIILCSVFAV--VT 299
 Db 241 KLPACSCSDCLSCPGPPPPPPPPKIVGLDAVF-----VIMAAVPLVGLV 288
 QY 300 ILLVG---PRVAPARDKSKMVDPKGTSLSDKLSFS-----THLLGQFQGW 344
 Db 289 VFLMGSLFTQSSMDNPFVD---GNDVSDEMPYSENDSEYFEKLGATHTTETLETFTKW 345
 QY 345 GTWVASWPLTIIVLSVIPVVALAAGLVFTELTTPVELWSAPNSQARSKAFHQHFGPF 404
 Db 346 GTYFASNPGTLIAGASLVVILGYGINFIEITDPVKLWASPNKSLREBREFDTKSPF 405
 QY 405 FRTNQVILTPAPRSSYRVYDLSLLGPNFSGIILDLLELLELQERLRLHQLVWSPBAQRN 464
 Db 406 YRLEQIIKAVNLPOIVHNT--SNGPYTPGVPDRBFLTKVLDLQEGIKEINA-----NG 458
 QY 465 ISLDICVAPLNDYSL--YDCCINSLLOYQONRTLLLTANQTLMOQTQSDWKDHP 522
 Db 459 TQLKDICTAPLSDGSEIDVSCVQVQSWGFGDDRERL----DDHEDENGFNVTYLDAL 514
 QY 523 LYCANAPLTFKGTALALSCMADYCAPVFPFLAIGY-----KGKDYSAEALIMTF 574
 Db 515 YDCISNPLY-----CLAPYGGVPDPAIAGGLFPPGDLQGTSTKPELANAILTF 564
 QY 575 SLNNYPAGDPRLAQAKLEAEFLEENRAF--QRRMAGMFQVTFATERSLEDEINRTAEDL 633
 Db 565 LVKNH--HNKTDLENALTWEKKFVEFMTNTYKNMSQYMDIAFTSERSIEDELNRESQSDV 623
 QY 634 PIFATSYIVIFLYISLALGSYSSMRVMDSKATLGLGVAVVLGVANMAAGPFSYLGTR 693
 Db 624 LTLVSYLIMFYIAISLGHVKEFKRVFIDSKITLIGGVIIIVLASVWSSVGVFGYGLP 683
 QY 694 SSLVILQVVPFLVLSVGDNIPIFVLEYQRLPRRPGEPREHIGRALGRVAPSMILCSLS 753
 Db 684 ATLIIIVEIPFLVAVGVDNIFILVQTHORDORKNETLEQQVGRILGKVGSMILTSLS 743
 QY 754 EACIFFLGAITPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKROEASRLDVCCCV 813
 Db 744 ESFCFFLGLSLDMPAVRALYAGVALIIDFLQLTCTVSLTFLDTKREERNMDICCFI 803
 QY 814 KPQELPPQCGGELLGFFQKAYAPFLHMITRQVLLLFLALFGVLSYVMCHISVGLDQ 873
 Db 804 KGKPKDSITSNELLYKFPSSVYVVPFLMKIVRASVMVIFPAWLCFSTAIAPRIDGLDQ 863
 QY 874 ELALPKDSYLLDYFLNRPFEVGAFFVFTTLGYNFSSEAGMAICSSAGNNFSFTOK 933
 Db 864 ELAMPQDSFVLHYFQSLNENLMPGVVFLVGLDLYATNNSQNLVCAQYCNDSVLQ 923
 QY 934 IOYATEFPEQSYLAIPASSWVDDFDIM--LTPSSCCRLYISGPNKDK--PCPSTVNSLNC 990
 Db 924 IYLASHNSQTYIARPASSWIDYFDMAAASCCY-----RKSGDFCPH--QDTSCL 976
 QY 991 KNCMSITWGS--VRPSVQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDQVLI---A 1046
 Db 977 R--CNITKNSLLRPEKEFVKYLPFLKXNPDPTCAKAGHAAYGAVRVYSSNSHERLNIEA 1034

```
QY 1047 SFEMAYHKLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAREVEPYTITN 1099
Db 1035 SYFMAYHTLKSSADYFLALESARKISANITQMLQRLMSNGVPWASALTVEVFPYSVYF 1094
QY 1100 VFYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLMLLSIVMLVDTVGFMAIW 1159
Db 1095 VFYEQYLTMWSDTLQSMGIVSLVIFVTVFLMGFDVHSAIVVITITMIVNLGLMYW 1154
QY 1160 DISYNAVSLINLVASGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAVAMT 1219
Db 1155 NISLNAVSLVNLVMAVIGSVFCSHLVHSFATSQVSDRAADSLSKMGSSIFSGITLT 1214
QY 1220 NLPGLVILGLAKAQLIOIFFFFRLNLTITLLGLLHGLVFLPVLTLSVVGDPVNPALAEQKR 1279
Db 1215 KPAGILVLAFAKSQIFQVFFRMYLGI VVIGAAHGLIFLPVLLSYIGAPVSNARLRYHSQ 1274
QY 1280 A-----EEAVALVM 1288
Db 1275 AAHEHETALAGIL 1287

RESULT 90
US-09-614-150-12003
; Sequence 12003, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-12003

Query Match 32.7%; Score 2262; DB 26; Length 1287;
Best Local Similarity 37.8%; Pred. No. 6.1e-189;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LIAVLLRLAQSEPYTHIQPGYCAPYDECGKNPELSGLMTLSNVSLSNTPARKITGD 69
Db 17 LIAAVLFTLIQSSKQD-----CWVYGVCTNN-DFSHS-----QNCYPNGTAKEMATD 62
QY 70 HLILKQICPRLYTGNTQACSAKOLVSLASLSITKALLTRCPACSNFVNLHCHNTC 129
Db 63 GLELLKKRCGLFLENSKFPCCQKNQVLLNKNVLAGNILLDRCPSCMENLVRHICQFTC 122
QY 130 SPNQSLFINVTRVAQLGAGOLPAVAYAEAFYQHSFAEQSYDSCSRVRVPAATAVAGTWC 189
Db 123 SPKQAEFMHVATQKNKGD-EYISSVDLHISTEYINKTKSCSQSVSPQTGQLAFDLMC 181
```

```
QY 190 GYVSALCNAORWLNFOGDTGNGIAPLDITFHLLEPGQAVGSGIOPLNGVARCNESQCD 249
Db 182 GAYSASRCNPTKWNFMGDATNPVVPQIYIQHBP-KSNNNFTPLNVTTPCNOAVSS 240
QY 250 DVATCSQDCDCAASC-----PAIARPQ-----ALDSTFYLGQMPGSLVLIIILCSFPAV-VT 299
Db 241 KLPACSCDCDLSQCPQPPPEPPPEPKIVGLDAVF-----VMAAVFLVGLV 288
QY 300 ILLVG----FRVAPARDKSKMVDPKKGTSLSKLSFS-----THLLGQFFQW 344
Db 289 VFLMGSEFLTGQSSMDONFQVD--GNDVSDEMPYSENDSYFEKLGAGHETFTFLETFYTKW 345
QY 345 GTWASWPLTTLVLVSVIPVALAAGLVFTELTTPDVELWASPNQARSEKAFHQHQPFP 404
Db 346 GTTFASNPGLTLIAGASLVILGYGINFIEITTPVKUWASPNQSRLEREFDFTKFEPF 405
QY 405 FRTNQVILTAPNRSYRYDSLLGLPKNPSGILDLDLLELLELQERLRLHQLWSPQAQRN 464
Db 406 YRLEQIIIKAVNLQIVHNT-SNGPYTFGVPDFREFLTKVLDLQBGIKEINA-----NG 458
QY 465 ISLQDICYAPLNPDNTSL--YDCCINSLQYFQNNRTLLLTANQTLMGQTSQVQDWKHF 522
Db 459 TQLKIDICYAPLSDGSEIDVSCVQSVQSIWGYFGDDRERL-----DDHDDNGFNVTYLDAL 514
QY 523 LYCANAPLTFKDGTAALALSCMADYGAPVPFPLAIGGY-----KGKDYSEAEALIMTF 574
Db 515 YDCISNPYL-----CLAPYGGPVDPAIALLGGFLPGDQLTGSKTFELANAILITF 564
QY 575 SLNNYPAGDPRLPAQAKLWEEAFLEEMRAF-QRRMAGMPQVTTAERSLEDEINRTABDL 633
Db 565 LVKNH-HNKTDLLENALTWEKKFVERMTNNTKNMSQYMDIAFTSERSIEDELNRESQSDV 623
QY 634 PIPATSYIVIFLYTSLALGSYSSWSRVWDSKATLGLGVAVVLGAVMAAMGFYGLGR 693
Db 624 LTIIVSYLIMFYIAISLGHVKFKEFVDSKITIGIGVIVLASVSSVGVFGYIGLP 683
QY 694 SSLVTLQVVPFLVLSVGADNIFIVLEYORLPRRPGEPREHIGALGRVAPSMLLCSLS 753
Db 684 ATLIIIVEIPFLVLAAGVDNIFILVQTHQROKRENETLEQQVGRILKVGSMILTSLS 743
QY 754 EACFFLGALTPMAVRTFALTSLGLAVIDLFLQLMSAFALLSLDSKQEAERLQVCCV 813
Db 744 ESPCFFLGSLDMPAVRAFALYAGVALLIIDFLQITCFVSLFTLDTKREENRMDICCFI 803
QY 814 KPQELPPPGQGBGLLGFQKAYAPFLHWTGRVLLFLALFGVLSYMSCHISVGLDQ 873
Db 804 KGKKPDSITSNEGILLYKFSSVYVFPMLKKIVRASVMVIFFAWLCSFAIAPRIDGLDQ 863
QY 874 ELALPKDSYLLDYFLFNRYFEVGPVYFVTLTGYNFSSEAGMNAICSSAGCNFSFTOK 933
Db 864 ELAMPQDSFVLHYFOSLNENLNGIPVTVLKGDLAYTNSSQNLVLCAGQYCNDDSVLTQ 923
QY 934 IQYATEFPPEQSYLALPASSWDDFDIDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNC 990
Db 924 IYLASRHSNQYIARPASSWIDDDYFDMAAASSCKY-----RKDSGDFCPH-QDTSCL 976
QY 991 KNCMSITMGS-VRPSVEQFHKLWPFLNDRPNKICPKGLAAYSIVNLTSQGVL--A 1046
Db 977 R--CNITKNSLLRPEKEFVKYLPFLKDNPDQTCAGAHAAYGAVRYSNSHERLNIEA 1034
QY 1047 SFEMAYHKLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAREVEPYTITN 1099
Db 1035 SYFMAYHTLKSSADYFLALESARKISANITQMLQRLMSNGVPWASALTVEVFPYSVYF 1094
QY 1100 VFYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLMLLSIVMLVDTVGFMAIW 1159
Db 1095 VFYEQYLTMWSDTLQSMGIVSLVIFVTVFLMGFDVHSAIVVITITMIVNLGLMYW 1154
QY 1160 DISYNAVSLINLVASGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAVAMT 1219
Db 1155 NISLNAVSLVNLVMAVIGSVFCSHLVHSFATSQVSDRAADSLSKMGSSIFSGITLT 1214
```

QY 1220 NLPGILVLGAKAQLIQIFPRRLNLLITLLGLLHGLVFLPVLISYVGPDPNPALEQR 1279
DB 1215 KPAIGLVAPAKSQIFQVFPYRMYLGIIVVGAHGLIFLPLVLLSYGAPVSNARLYHSQ 1274
QY 1280 A-----EEAVALM 1288
DB 1275 AAHEHETALAGIL 1287

RESULT 91
US-09-614-150A-12003
; Sequence 12003, Application US/09614150A
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150A
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PR1
; ORGANISM: DROSOPHILA
US-09-614-150A-12003

Query Match 32.7%; Score 2262; DB 26; Length 1287;
Best Local Similarity 37.8%; Pred. No. 6.1e-189;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LIAVLLRLAQSEPTTHIQPGYCAFYDECGKNPELSGLMTLSNVCLSNTPARKITGD 69
DB 17 LIAVLLRLAQSEPTTHIQPGYCAFYDECGKNPELSGLMTLSNVCLSNTPARKITGD 62
QY 70 HILLOKICPRLYTGNTQACSAKQLVSLASLITKALLTRCPACSNFNLHCHNTC 129
DB 63 GLELLKRCGGLLENSKFCDDKQVLLNKNVLAAGNILDRCPCSMENLVRIHQFTC 122
QY 130 SPQSLFINVTRVAQLGACQLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATLAVGTCM 189
DB 123 SPQAEFMHVATQNNKGD-BYISSVDLHISTEYINKTKSCSQSVSPQGLADFLMC 181
QY 190 GYVGSALCAORLWLPQGTGNGLAPLDITPHLLBPQAVGSGIOPLNEGVARNCSQGD 249
DB 182 GAYSASRCNPTKWFNMGDAPYVFPQIYIQHEP-KSNSNNFTPLNVTTPCNQAVSS 240
QY 250 DVATSCQDCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSFVAV-VT 299
DB 241 KLPACSCSDCLSCQGPPEPRPPEPKIVGLDAYF-----VIMAAVFLVGL 288
QY 300 ILLVG---PRVAPARDKSNVDPKKGTSLSDKLSFS-----THLLGQFFQGM 344
DB 289 VPLMGSLFTQSSMDNFDQVD---GNDVSDMPYSENDSYPEKLGHAHTETLETFTFTK 345

RESULT 92

US-11-097-143-12003

; Sequence 12003, Application US/11097143

; GENERAL INFORMATION:

QY 345 GTWVASWPELTILVLVSVIPVWALAAAGLVFTTELTTDPVVELMSAPNSQARSSEKAFHDHQFGB 404
DB 346 GTYPASNPGCLTLIAGASLVILGYGINFTEITITDVPVKLWASPNKSRLEBEPFDTKFSDF 405
QY 405 PRTNQVILTAPNRSSRYRDSLLGLPKNPSGILDLLELLELQERLRLHQVWSPQAQRN 464
DB 406 YLEQIIIIKAVNLPOIVHNT-SNGPYTTPGVDFRFLTKVLDLQBGKEINA-----NG 458
QY 465 ISLODICYAPLNPNTSL--YDCCINSLLQYQNNRTLLLTANQTLQCTSQVOWKDH 522
DB 459 TQLKIDICYAPLSDGSEIDVSCVQSIWGYGDDRERL-----DHDDEDGNFVTVLDAL 514
QY 523 LYCANAPLTFKDTALALSCMADYGAPVPPFLAIGY-----KGKDYSAEALIMTF 574
DB 515 YDCISNPYL-----CLAPYGGPVDPAIALGGLFPPGDLTGCTSTKSTFELANAILTF 564
QY 575 SLNYPAGDPRLAQAKLMEAEFLEWRAP-QREMGMEQVTFEARSLEDEINRTTAEOL 633
DB 565 LVKNH-HNKTDLLENALTWEKKFVEFTMYTKNNMSQYMDIAFTSEKSIDELNRSQSDV 623
QY 634 PIFATSYIVIFLYISIALGSSYSSMRVMDSKATLGLGGVAVVLGAVMAAMGFFSYLGR 693
DB 624 LTILVSYLIMFYIAISLGHVKEFKRVFIDSKITLIGGVIIIVLASVSSVGVGIGLP 683
QY 694 SSLVILQVVPFLVSVGADNIPFVLEYQRLPRRPGEPREVNHIGRALGRVAPSMILCSLS 753
DB 684 ATLIIIVEIPFLVAVGVNIFILVQTHQDQKQNETLEQQVGRILKGVGFSMLLTSL 743
QY 754 EACIPELGNLTPMPAVRTEALTSGLAVILDELLQMSAFVALLSLSKROEASRLDVCCCV 813
DB 744 ESFCFLGGLSDMPAVRAPALYAGVALLIDFLQLITCFVSLFTLDTKRRENRMDCCFI 803
QY 814 KPQELPPPQCGEGLLGPFQKAYAPFLHWTGVRVLLILFLALFGVLSYSMCHISVGLDQ 873
DB 804 KGKPDSTISNEGILYKFSVVYVFLMKKIVRASVMVIFPAWLCFSAIAIRIDIGLQ 863
QY 874 ELALPKDSYLDYFLFLANRYPEVGAPVYVTVTLGYNFSSEAGMNAICSSAGCNFSTQK 933
DB 864 ELAMPQDSFVLHYFQSLNENLNIGPPVYFVLKGLDAYTNSQNLVCAQYCNDDSVLTQ 923
QY 934 IQYATEPPEQSYLAIPASSWDDPDIDW-LTPSSCCELYISGNKDK--FCPSTVNSLCL 990
DB 924 IYLSRHSNQTYIARPPASSWIDDDYFDMAAASSCKY-----RKDSGDFCPH--QDTSCL 976
QY 991 KNCMSITWGS--VRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQDQVL---A 1046
DB 977 R--CNITKNSLLRPEKEFVKYLPFLKONPDCTAKAGHAAYGCAVRYNSHERLNIEA 1034
QY 1047 SRFMAYKPLKNSODYTEALRAARELANITADLR-----KVP-GTDPAPFVFPYITN 1099
DB 1035 SYFMAYHTLKSSADYFLALBSARKISANITQMLQGLMSNGVPMASALTVEVPVSVFY 1094
QY 1100 VPYEQYLILPEGLPMLSLCLVPTTAVSCLLGLDLRGLLNLISIVMLVDTVGFMAW 1159
DB 1095 VPYEQYLTWMSDTLQSMGIVSLISFVTVFLMGFDVHSALVVVITITMIVNGLGLMYM 1154
QY 1160 DISYNAVSLINIVSAGMSVBFVSHITRSPALSTRPTWLERAKEATISGMSVAFAGVAMT 1219
DB 1155 NISLNAVSLVNLVMAVGSVEPCSHLVSPATSKVSQIDRAANDLSKMGSSIFSGIUT 1214
QY 1220 NLPGILVLGAKAQLIQIFPRRLNLLITLLGLLHGLVFLPVLISYVGPDPNPALEQR 1279
DB 1215 KPAIGLVAPAKSQIFQVFPYRMYLGIIVVGAHGLIFLPLVLLSYGAPVSNARLYHSQ 1274
QY 1280 A-----EEAVALM 1288
DB 1275 AAHEHETALAGIL 1287

```
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-12003

Query Match          32.7%; Score 2262; DB 40; Length 1287;
Best Local Similarity 37.8%; Pred. No. 6.1e-189;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LIAALLRLAQSEPTTHIQPGYCAFYDECGNPNLSGLMTLSNVSCLSNTPARKITGD 69
DB 17 LIAAVLFTLIQSKQD-----CWVYGVCTN-DFSHS-----QNCFYNGTAKEMATD 62
QY 70 HUILOKICPRLYTGNTQACSAKQLVSLASLITKALLTRCPACSDNFVNLHCHNTC 129
DB 63 GLELLKRCRGFLLENSKFCCKQKQVELLNKNVELAGNILDRCPCSMENLVRIHCQPTC 122
QY 130 SPQSLFINVTRVAQGLQAPVAVAYEAFYQHSFAEQSYDSCSVRVPAATLAVGTC 189
DB 123 SPQAEFMHVATQNNKGD-BYISSVDLHISTEYINKTYKSCSVSPVQGTQLAFDLMC 181
QY 190 GYVSALCNAORLWLFQDGTGNGLAPLDITFHLBPQAGVSGIOPLNEGVARCNESQGD 249
DB 182 GAYSASRCNPTKWFNMGDATTNPPYFQIYIQHEP-KSNSNNFTPLNVTTPVCNQAVSS 240
QY 250 DVATSCQDCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSVPFV-VT 299
DB 241 KLPACSCSDCLSCQGPPEPRPFPKIVGLDAYF-----VIMAAVFLVGLV 288
QY 300 ILLVAG---PRVAPARDKSNVDPKGTSLSDKLSFS-----THLLGFOFGW 344
DB 289 VFIMGSLFTQSSMDNDNQVD---GNVSDMPSEYSENDSYEKIGAHTETPLETFTFTKW 345
QY 345 GTWVASWPLTILVSLVIPVALAAGLVFTLITDPVELMSAPNSQARSEKAFDQHGPF 404
DB 346 GTYFASNPGCLTIAGASLVILGYGINFIEITTDVPKLWASPNKSLRLEBFDDTKFSPF 405
QY 405 FRTNQVILTAPNRSSRYRYSLLGNPNFSGIILDLLELLELQERLRLHQLWSPSAQRN 464
DB 406 YRLEQIIKANVLPQIVHNT-SNGPYTFGVPDREFLTAKVLDQEGIKEINA-----NG 458
QY 465 ISLQDICVAPLNDTSL--YDCCINSILQYFQNNRTLLLTANQTLMGQTSQVMDKHP 522
DB 459 TQKDLICVAPLSDGSEIDVSCVQVSIWGYGDDRERL-----DDHEDNGFNVTYLDAL 514
QY 523 LYCANAPLTFKDTALALSCMADYGAPVPFPLAIGGY-----KGKDYSEAEALIMTF 574
```

```
DB 515 YDCISNPFY-----CLAPYGGPVDPAIALGGFLPPGQDLTGSTKPFELANAILTF 564
QY 575 SLNNYPAGDPRLAQAKLWEEAFLEEMRAF-QRRMAGMFQVTTFTABRSLEDEINRTTAEDEL 633
DB 565 LVKNH-HNKTDLLENALTWEKKFVEFTWNTYTKNNMSQYMDIAFTSERSIEDELNRSQSDV 623
QY 634 PIPATSYIVIFLYISLALGSYSSRSRVWVDSKATILGGGAVVLGAVMAAGFFSYLGR 693
DB 624 LTILVSLIMPMYTAISLGHVKFKRFIDSKITLIGIGGVITVLASVSSVGVFGYIGLP 683
QY 694 SSVLTVQVVPFLVLSVGADNIFIFVLEYQRIQPRPGEPREVHIGRALGRVAPSMLLCSLS 753
DB 684 ATLIIVEVIFELVAVGVDNIFILVQTHORQORPENETLEQQVGRILKGVPSMLLTSL 743
QY 754 EACFFFLGALTPMPAVRTFALTGSLAVILDLQMSAFVALLSDSKQOASRLDVCCV 813
DB 744 ESFCFFLGLGSDMPAVRAFALYAGVALIIDFLQLITCFVSLFTLDTKREERNMDICCFI 803
QY 814 KPQELPPPGQEGLLLGFFQKAYAPPELLHWTGRGVLLFLALFGVLSYSMCHISVGLDQ 873
DB 804 KGKPDSTISNEGGLLYKFFSSVYVFFLMKKIVRASVMVIFFAWLCSIAIAPRIDIGDQ 863
QY 874 ELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLTGYNFSSEAGMNAICSSAGCNFSFTQ 933
DB 864 ELAMPQDSFVLHYFQSLNENLNIQPPVYFVLKGLDAYTNSSDQNLVCAGQYCNDDSVLTQ 923
QY 934 IOYATEPPEQSYLAIPASSWVDDFDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNL 990
DB 924 IYLASRHSNQTIIARPASSWIDDFDMAAASCCY-----RKDSGDFCPH--QDTSCL 976
QY 991 KNCMSITMGS--VRPSVEQFHKVLPWFLNDRPNIKCPKGLAAYSTVNLTSQGVL--A 1046
DB 977 R--CNITKNSLURPEKEFVKILPFLKDNPDCTCAKAGHAAYGAVRYSHSHERLINEA 1034
QY 1047 SRFMAYHKLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAPFEPFYITN 1099
DB 1035 SYFMAYHTLKSSADYFLALESARKISANIITQMLQRLMSNGVPMASALTVEVFPYSVFY 1094
QY 1100 VFYQYLTILPEGLFMLSLCLVPTFVAVSCLLGLDLRLSGLLNLSIVMLVDTVGFMALW 1159
DB 1095 VFYEQYLTMTWSDTLQSMGISVLSIFVTVFLMGDFVHSALVVVITITMIVVNLGLMYYW 1154
QY 1160 DISYNAVSLINLAVGMSVEFVSHITRSPFAISTKPTWLERAKETISMGSAVPAGVAMT 1219
DB 1155 NISLNAVSLVNLVAVGISVEFCSHLVHSFATSKSVSQIDRAADSLSKMGSIFSGITLT 1214
QY 1220 NLPGLVLGLAKAQLIQIFFFLNLILTLGLLHGLVFLPVILSYVGPVDPNPALEQKR 1279
DB 1215 KPAIGILVAFAKSQIFQVYFPMYGLIIVIGAAHGLIFLPVLLSYIGAVSNARLRYHSQ 1274
QY 1280 A-----EEAFAVM 1288
DB 1275 AAABEHETALAGIL 1287

RESULT 93
US-60-229-511-274
; Sequence 274, Application US/60229511
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PHASE I PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000767
; CURRENT APPLICATION NUMBER: US/60/229,511
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: HUMAN
```

US-60-229-511-274

Query Match 30.2%; Score 2087; DB 46; Length 1525;
Best Local Similarity 32.3%; Pred. No. 2.3e-173;
Matches 515; Conservative 237; Mismatches 461; Indels 382; Gaps 38;

QY 33 CAPYDECG-----KNPELSGLMTLSNVSCLSNTPARKITGDHLLILQKICPRLYTGP 85
DB 9 CWYGECEGIAYGDKRYNCEYSG-----PPKPLPKDG-----YDLSAGFFFG- 49

QY 86 NTOACCSAKOLVSLBASLSITKALLTRCPACSDNFVNLHCHNTCSNQSLFNVTR---- 141
DB 50 NVSLCCDVRLQTLKDNQLPQLFSLRCPSCFNLLNLFCELTCSRQSQFLNVTATEDY 109

QY 142 VAQLGAGQIPAVVAYEAFYQHSPABOSYDSCSRVRVPAATLAVGTMCVYGVSALCNAQR 201
DB 110 VDPVNTQTKNVKELQYVVGQSFANAMYNACRDEAPSSNDKALGLLCKGDADA-CNATN 168

QY 202 WLNFGQDTGNGLAPLDIT-----PHLEPGQAVGSGIOPNEGVARNCSQGGDVATC 254
DB 169 WTEYMFNKGNGQAPFTTIPVFSDFPVH-----GMEPMNNAKGCDSVDEVTAPC 218

QY 255 SCODCAASCAPATARPQ-----ALDSTFVLGQMPGSLVLIILICSPFVAVTILL--- 302
DB 219 SCQDCSIVCGPKPQPPPPAPWTILGLDAMYIMMITYMAFLLVFFGAPFAVWCYKRYF 278

QY 303 -----VGRFV-----APARD-----KSKWV-----DP 319
DB 279 VSEYTPIDSNIAFVSNASDKVGLQWVRWCASAEATATKADMVLAFAVLPSSQGGTAEP 338

QY 320 KIGTSLSDKLSSTHTLLQPFQGGTGWASWPLTILVSVIPVVALAAGLVPTBLTTPD 379
DB 339 RE-ASCCDPVSAFAFGCLRLRTRWGSFCVRNPGCVIPFSLVFITACSSGLVFVRVTNP 397

QY 380 VELWAPNSQARSEKAFDHOHGPPRTNOVILTAPNRSSYDYDSLLGPK-NFSGILD 438
DB 398 VDLWAPSQARLEKEYPDQHGPPRTTEQLIIRAPLTDKHYQYPPSGADVPPGPPLDI 457

QY 439 DILLLELLEQLRLHLQWSPQAQNISLQDICYAPLNPNDTSLYDCCINSLLQYFQNNR 498
DB 458 QILHQVLDLQIAIEN--ITASVDNETVTLQDILCLAPLSYNT--NCTILSVLNFQNSH 512

QY 499 TLLLLTANQTLMGQTSQVDMDKHFLY----- 524
DB 513 SVLDDHKKGGDDF--VIADYHTHFLYCVRVGDGAEFWISAVDEGTQLLAAGHLTSLMSP 569

QY 525 -----CAN----- 527

DB 570 IHSDDGWICRSPSAPGVVADGLSENAGSLSTGQLSVCTALPVASCLGHCANISTYCSTC 629

QY 528 -----APLT 531

DB 630 FLFLCFRQTFHFAQAI PVKQLMTFLSRKSLIVPQCLQNKGNPNYLTGTESFYFAFCPLL 689

QY 532 FK-----DGTALAL-----SCMADYGAP 549
DB 690 PPDWPPPSGAPLRHAPCAIDGTGLYPASCVPVILSSQAPASLNDTSLIHDPCLTGFGP 749

QY 550 VPPFLAIGYKGYDSEBALMTWLSLNNYPAGDPLAQKLWEAFLEEMKAFORRMAG 609
DB 750 VPPWLVLGGYDQNNATALTIVTPVNNYNDTEKLRQAQWEKEFINFVKNYKN---P 806

QY 610 MFQVFTFAERSLEDEINRTAEDLPIFATSYIVIFLXISLALGSSYSSMSRVMDSKATIG 669
DB 807 NLTISFTAERSLEDEINRESDDVTTVISYAIMFLYISLAUGHKSKRLLVDSKVSIG 866

QY 670 LGGVAVLGVANMAAGFFSYLGIRSLVLOVVPFLVSVGADNIFIFVLEYQRLPRRP 729
DB 867 IAGILVLSVACSLGVFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAVQ----- 919

QY 730 EPREVHIGALGRVAPSMMLCSLSBAICPFL-----GALTTPMVAVNTFALTSL 778
DB 920 -----VYPHILTGLSICS-SPRDCWEILGSAFVSVTCGALSVMVAHVHTFSLFAGL 969

RESULT 94

US-60-212-657-129

; Sequence 129, Application US/60212657

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

; FILE REFERENCE: CL000672

; CURRENT FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 303

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 129

; LENGTH: 1529

; TYPE: PRT

; ORGANISM: HUMAN

US-60-212-657-129

Query Match 30.1%; Score 2083; DB 46; Length 1529;

Best Local Similarity 32.3%; Pred. No. 5.2e-173;

Matches 517; Conservative 236; Mismatches 457; Indels 392; Gaps 39;


```
QY 142 VAQLGAGQLPAVAVAEAFQHSFASQSDSCSRVRVPAATAVAGTMCVGYGSALCNAQR 201
Db 110 VDPVTNQTNTKVKELQYVVGQSFANAMYNACRDVEAPNSNDKALGLCCGKORADA-CNATN 168
QY 202 WLNFOQDGTNGIAPLDIT-----PHLLPQAVGSGIQIPLNEGVARCNESQGDVATC 254
Db 169 WIEYFNKDNQGAFTIITPVFSDFPH-----GMEPNNAATKCDSDSDVETAPC 218
QY 255 SQDCNASCAPAIARQ-----ALDSTFYLGMPGSLVLIILCSVFAVVTILL--- 302
Db 219 SQDCSIVGKRPQPPPPAPWTILGLDAMYIMMITYMAFTLLVFFGAFFAVWCYKRYF 278
QY 303 -----VGRV-----APARD-----KSKWV-----DP 319
Db 279 VSEYTPIDSNIAFSVNASKVGLQWVRWCASAEQDTAATKADMLAFVALPSSQGGTAEP 338
QY 320 KKGTSLSLDSFTHTLLQFFQGGTGWASWPLTILVLSVTPVVALAAGLVFTLLTDP 379
Db 339 RE-ASCCDPVSAFEGCLRLTRWGSFCVRNPGCVIPFSLVFIACSSGLVFRVVTNP 397
QY 380 VELWAPNSQARSEKAFHQHGFRTNQVILTAPNRSYRYDSLLGPKNFSGILDL- 438
Db 398 VDLWAPSQARLEKEYFQHFQGFRTQLIIRAPLTDKHIQPYL---RELMPDLDR 454
QY 439 -----DLLLELLELLELRLHQLVMSPEAQRNISILODICYAPINPDNTSLYDCCINSLL 491
Db 455 LYRYCTSLLLKVLQIAIEN--ITASYDNBTVTIQLQICLAPLSPYNT---NCTILSVL 509
QY 492 QYFQNNRTLLLTANQLMGQTSQVDWKDHFLY----- 524
Db 510 NYFQNSHSLVDHKGGDDFF--VYADYHTHFLYCYVRVDGADEPWISAVDEGTQLLAAGHL 566
QY 525 -----CAN- 527
Db 567 TSLMPSPHSDGWVICRSPSAPQVQVADGSLGNAGSLSTGQLSVCTALPVASCLGHCANI 626
QY 528 ----- 527
Db 627 STYCTCFPLCRQTFTHATQAI FVKQLMTFLSRKSLIVPQCLQKNPNYLTGTESFYF 686
QY 528 ---APLTFK-----DGTALAL-----SC 542
Db 687 AFVCPLLFPDPPPGAPLRHAPCAIDGTGLVLPASCVPILSSQAPASLNDTSLLDHPC 746
QY 543 MADYGAPVFPFLAIGYKGDYSEBALIMTSLNYPAGDPRPQAOKLWEAPLEEMRA 602
Db 747 LGTFGQVFPFLVGLGDDQNNATALTITFPVNNYNDTEKLQRAQAWKEFEINFKVN 806
QY 603 FORMMAGMFQVFTFAERSLEINRTAEDLPFATSYIVIFLYISLALGSSYSNRSVMV 662
Db 807 YKN---PNLTISFTAERSLEDELRNEDSDVFTVVISVAIMFLYISLAUGHYKSCRRLLV 863
QY 663 DSKATILGLGGVAVLGNVMAAGFFSYLGIKSLVILQVVPFLVLSVAGADNIFIFVLEYQ 722
Db 864 DSKVSLIGIAGILVLSVACSLGVSYIGLPLTIVIEVIFPLVLAGVDNIFILVQAYQ 923
QY 723 RLPRRGPPEVHIGRALGRVAPSMLLCSLSBAICPFL-----GALTPMPAVRT 771
Db 924 -----VYPHILTGLSICS-SFRDCEWLIGSAFVSVTCGALSVMPAVHT 966
QY 772 FALTSLGLAVILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPGQ-GRGLLIG 830
Db 967 FSLFAGLAVIFDLQITCFVSLGLLDIIRQENRLDIFCCVRGAEDGTGVSQASSCLFR 1026
QY 831 FFKAYAPFLHWTIRGVVLLFLALFGVSLYSMSCHISVGLDQELAPKDSYLLDYFLFL 890
Db 1027 FFKNSYSPLLXKDWMBPIVIAIFVGVLSPFSI AVLNKVDIGLQSLSPMDSDSYMDVFKSI 1086
QY 891 NRYFEVGAPVYFTTILGYNFSSEAGNAICSSAGCNFNFTQIKIYATFEPQSYLAIPA 950
Db 1087 SQYLHAGPPVYFVLEEGHDYTSKQNMVCGMGNCNDSLVOQIIFNAALDNYTRIGFAP 1146
QY 951 SSWVDDFIDLWLP-SSCRLXYISGPNKDKFCPSTVNSLNCNCKMSIT-MGSVRPSVEQF 1008
```

```
Db 1147 SSWIDDDYDWKPOSSCCRV---DNITDOFCNASVVDPAQVR-CRPLTPEGKORQGGDF 1202
QY 1009 HKYLPWFLNDRPNIKCPKGGLAAYSTVN--LTSDDQVQLASRFMAHYHKPLXNSQDYTRAL 1066
Db 1203 MRFLPMFLSDNPNPKCGKGGHAYSSAVNILLGHGTRVGATVFTYHTVLOTSAFDIAL 1262
QY 1067 RAARELANIATADLRKVPDTPAPEVPPYTTINPYEQYLITLPEGLFMLSCLVPTTAV 1126
Db 1263 KKAHLIASNV7-ETMGINGS--AYRVFPYSFYFYEQYLTIIDTTIENLGVSLGAIFLV 1319
QY 1127 SCLLGLDLRGLGLNLLSIVMILVDTVGFMAWDISYNAVSLINLV----- 1172
Db 1320 TWVLGCELMSAVIMCATIAMVLMFVGMVLMGLISLNAVSLVNLVWGLQBEESHPKSA 1379
QY 1173 --SAVGMSEVFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLA 1230
Db 1380 LPQSGISVEFCSHITRAFTVSMKGSRVERAEALAHMGSSVFSGITLTKEGIVVLAPA 1439
QY 1231 KAQLIQIPFFRLNLLITLLGLHGLVFLPVLISYVGPDPNPALEEQKRAEBAVAAMVA 1290
Db 1440 KSOIQIPIFYFMYLAVILLGATHGLIFLPVLSYI-----ALSQGILTCLURV--- 1488
QY 1291 SCPNHSRVSADNIY----VNHSFEGSIKGAISNPLNN 1328
Db 1489 -----GRFTTGRVLHRQGVHEHRMVPITGCLAAALNAVAVN 1524

RESULT 95
US-60-191-637-12035
; Sequence 12035, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12035
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-12035

Query Match          29.5%; Score 2036.5; DB 45; Length 1157;
Beat Local Similarity 36.5%; Pred. No. 4.2e-169;
Matches 457; Conservative 227; Mismatches 392; Indels 175; Gaps 27;

QY 92 SAKQLVLSLEASLSITKALLTRCPACSDNFVNLHCHNTCSPNQSLFINTRVLAQLGAGQLP 151
Db 28 SSKQVELLNKNVLAGNLDRCPSCMENLVRIHCQFTCSPKQAEPMHVVATQKNKKGD-E 86
QY 152 AVAYAEAFYQHSFAEQSYDSCSRVRVPAATAVAGTMCVGYGSALCNAQRWLNFOQDGTN 211
Db 87 YISSVDLHISTEYINKTKYCSQSKLPA----- 114
QY 212 GLAPLIDITFHLLEPQAVSGSQIPLNEGVARCNESQGDVATCSDCAASC-----PATA 267
Db 115 -----CSCSDCLSLCPQGPPEPP 132
QY 268 RPQ-----ALDSTFYLGMPGSLVLIILCSVFAV-VTILLVG-----FRVAPARDKSKMVD 318
Db 133 RPEPKIVGLDAYF-----VIMAAVFLVGLVFLMGSLFTQGSMDNDNFQVD 180
QY 319 PKKPTSLSDKLSFS-----THTLLGQRFQGGTGWASWPLTILVLSVIVVAL 366
Db 181 ---GNDVSEDEMPYSENDYFKEKLGHAHTETLETFTFKMGTTYPASNPGLTLIAGALSIVL 237
QY 367 AAGLVFTSLTTPVELWAPNSQARSEKAFHQHGFRTNQVILTAPNRSYRYDSLL 426
Db 1147 SSWIDDDYDWKPOSSCCRV---DNITDOFCNASVVDPAQVR-CRPLTPEGKORQGGDF 1202
QY 1009 HKYLPWFLNDRPNIKCPKGGLAAYSTVN--LTSDDQVQLASRFMAHYHKPLXNSQDYTRAL 1066
Db 1203 MRFLPMFLSDNPNPKCGKGGHAYSSAVNILLGHGTRVGATVFTYHTVLOTSAFDIAL 1262
QY 1067 RAARELANIATADLRKVPDTPAPEVPPYTTINPYEQYLITLPEGLFMLSCLVPTTAV 1126
Db 1263 KKAHLIASNV7-ETMGINGS--AYRVFPYSFYFYEQYLTIIDTTIENLGVSLGAIFLV 1319
QY 1127 SCLLGLDLRGLGLNLLSIVMILVDTVGFMAWDISYNAVSLINLV----- 1172
Db 1320 TWVLGCELMSAVIMCATIAMVLMFVGMVLMGLISLNAVSLVNLVWGLQBEESHPKSA 1379
QY 1173 --SAVGMSEVFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLA 1230
Db 1380 LPQSGISVEFCSHITRAFTVSMKGSRVERAEALAHMGSSVFSGITLTKEGIVVLAPA 1439
QY 1231 KAQLIQIPFFRLNLLITLLGLHGLVFLPVLISYVGPDPNPALEEQKRAEBAVAAMVA 1290
Db 1440 KSOIQIPIFYFMYLAVILLGATHGLIFLPVLSYI-----ALSQGILTCLURV--- 1488
QY 1291 SCPNHSRVSADNIY----VNHSFEGSIKGAISNPLNN 1328
Db 1489 -----GRFTTGRVLHRQGVHEHRMVPITGCLAAALNAVAVN 1524
```


Db 238 GYGINFIEITDVKLWASPNKSRRLERBFDTKFSFFTRLEQIIKAVNLPOIVHNT-S 296
QY 427 LGPNFSGILDLLLELLELQRLHRLQVMSPEAQRLNLSQDICYAPLNPNTSL--YD 484
Db 297 NGPYTGPFDREFLKVLQDQGIKEINA-----NGTQKXDCIYAPLSDGSEIDVQ 350
QY 485 CCINSLQVFNQNTLLLTANQTLMGQTSQVQWQKHFLYCANAPFTFDGTALALSCMA 544
Db 351 CVVQSIWGWFGDRLR-----DDHEDNGFNVTYLDALYDCISNPLY-----CLA 396
QY 545 DYGAPEPPLAIGGY-----KGDYSEAEALIMTFSLNNYPACDPRLAQAKLWEAP 596
Db 397 PYGGPVDPAALAGFLPPGDQLTGSKYFELANAILITFLVKNH-HNKTDLENALTWEKFP 455
QY 597 LEMRAAF-QRRMAGMQVFTFAERSLEDEINRTAEDLPFATSIYIVIFLISLALGSYS 655
Db 456 VERMTYTKNNMSQYMDIAFTSERSIEDELNRESQSDVLTILVSYLIMFYAISLGHVK 515
QY 656 SWSRVVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSLVLQVLPFLVLSVGDADNIF 715
Db 516 EFKRVFIDSKITLIGIGVVIIVLASVSVGVFGYIGLPATLIIVEPIPLVAVGVNDIF 575
QY 716 IFVLEYORLPREGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTFALT 775
Db 576 ILVQTHORQPKNETLEQQVGRILKVGPSMLLTLSSESFCFFLGLSDMPAVRAFALY 635
QY 776 SGLAVILD FLOWMSAFVALLSDSKRQZASRLDVCCKVPQELPPPGQCGEGLLGFQKA 835
Db 636 AGVALLIIDLQITCFVSLFTLTKREENRMDICCFIKGKPDSTITSNEGLLYKFSV 695
QY 836 YAPFLHWTGRVVLVLLFALFGVLSYMSCHISVGLDQBLALPKDSYLLDYFLNRYPE 895
Db 696 YVPFLMKKIVRASVMVIFPAWLCSIAIAPRIDIGLDQBLAMPQDSFVLHYFQSLNENL 755
QY 896 VGAPYVFTVTLGVNFSSEAGMNAICSSACNNPSPFKIQYATEPEEQSYLAIPASSWD 955
Db 756 IGPVYVFLKGLDAYTNSDQNLVCAQGYCNDSDVLTQIYLASHNSQTYIARPASSWD 815
QY 956 DFTDW-LTPSSCRLYISGNKDK--FCPSTVNSLNCNKMCSITWGS-VRPSVEQFHY 1011
Db 816 DFDWAAASCCY-----RKDSGDFCPH-QDTSCLR-CNITKNSLLRBEKEFVKY 866
QY 1012 LPWFLNDRPNIKCPKGLLAAYSTSVNLTSQGVY---ASRFMAHYHKLKNSQDYTBALRA 1068
Db 867 LPPFLKNDPDDTCAKAGHAAGVAVRYSHERLNTEASYMAYHTILKSSADYPLALES 926
QY 1069 ARELANIYADLR-----KVP-GTDPAPFVPPYTTNIFYEQYLTILPEGLFMLSCLV 1121
Db 927 ARKISANITQMLQGRLSNGVPMASALTVEVPYSYFYEQYLTWMSDYLQSMGISVL 986
QY 1122 PTFAVSCLLGLDLRGLLNLISVIMLVDTVGFMAWLDISYNAVSLINLVASGMSVRF 1181
Db 987 SIFVTFVLMGFDVHSALVVITIMVNLGGLMTYWNISLNAVSLNVLWAVGSLVRF 1046
QY 1182 VSHITRSPAIKPTWLBRAKATI SMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFR 1241
Db 1047 CSHLVHSFATSVSVS QIDRAADSLSKWSSIFSGITLTKFAGILVLAPAKSQIFQVYFR 1106
QY 1242 LMLLITLLGLLHGLVFLVLSVGVDPVNPALAEOKRA-----EEAVAAM 1288
Db 1107 MYLGI VVIGAAHGLIFLPVLLSYIGAPVSNARLRYHSQAAAEHETALAGIL 1157

RESULT 96

US-10-449-902-53619
; Sequence 53619, Application US/10449902
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Research.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 53619
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-53619

Query Match 27.5%; Score 1898.5; DB 34; Length 1279;
Best Local Similarity 34.0%; Pred. No. 7.7e-157;
Matches 450; Conservative 238; Mismatches 527; Indels 107; Gaps 31;

QY 3 EAGLGRWLLMALLRLLAQSEPYTTTHQPGYCAFYDECGRKNPELSGSLMTLSNVCSLNTNP 62
Db 10 EANL---IFTAFVLLASLISQSAQAQ---CVMYGQCAKNDR-----GQPLNCGYNGP 56
QY 63 ARKITGDHLLILKICPLRY-TGPNTQACCSAKQLVLSLEASLSITKALLTRCPACSDNFV 121
Db 57 PKPLDDEGLSILQKRCPHLVDSGDRATTCDDAAMLKBLDNVNAKVAGLFQRCPCSRNLL 116
QY 122 NLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAVAEFYQHGSFAEQSYDSCSRVRVPAAA 181
Db 117 GYWCMTCAPDOSLPMFVTKFGVGSSKV--VDANVNIHIDQEVLOATPFCCKGVWVWSSG 174
QY 182 TLAVGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLPEQAVSGIGIQLNEGVA 241
Db 175 KEADMWACGAHQASCTPQRFDFMGDPSPFPFKMN---ITSSTGTDFDMDIPRRPKFV 231
QY 242 RCNESQGDVATCSCODCAASC-ATARQALDSTFYLGOMPSLSVLIILCSVFAV 297
Db 232 SCSESFDNSSKPCVDCCEKAPKTDGSLLPPEKLNKWSFLGL--NLIVIIDIGISAALI 289
QY 298 VTILLVGRVAPAR----DKSKWVDPKGTSLSDKLSFSFTHLLGQFFQGWGTWASWPL 353
Db 290 VAALLWAFAGQKWLDEITPENK----LEKIGDSWSRFLERSFWLGTACAAHPV 345
QY 354 TILVLSVIPVVALAAGLVFTELTPDVELWSPNSQARSEKAPHDQHPGFRTNQVILT 413
Db 346 WVLFISSWIIAGLSYGVQFLVTTDPVQIWAAPNSLVROEKTFTDFPFPFVRAEQMFIK 405
QY 414 ANRSRYRYSLLGPKNPSGILDLLELLELQRLHRLQVMSPEAQRLNLSQDICYA 473
Db 406 AVGLDEVKHOT-PNGEITFGPVYNKEFLALRLDLMKITEI-----GKAEGAGLETVCILA 459
QY 474 P-----LNPDNLSYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVQWQKHFLYCAN 528
Db 460 PLASDFLGPVTS--QCTVFNIWGFQNSLEFP-----EKGRNYLQDINKCLR- 506
QY 529 PLTFKDGDTALALSCMADYGAFFPFLAIGGY--KKG-----DYSEABALIMTFPSLNNY 579
Db 507 -----NAYDPDCLASYGPAEPPIVGDYLEKQSVDRVDNADPKKARAAIILTFVLRS 559
QY 580 PAGDPRLAQAKLWEAFLEEMRAFQ-RMAGMQVFTFAERSLEDEINRTAEDLPFAT 638
Db 560 LNPD-EIEKAMKEKLYYDFMQNWTNTEKPDPMDFMDVAFSSERSVQDSEIERVSKSEVTVVI 618
QY 639 SYVIFLYISLALGSYSWSRVVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSLVI 698
Db 619 SYTVMFLYISLGRFTSWRTYLSKMTLGGVWVWVSVVCAVGFGYIGTPTMLT 678
QY 699 LQVPEFLVLSGADNIFIFVLEYQRLPRPG-----EPREVHIGRALGRVAPSMLLCSLS 753
Db 679 VEVIPLVAVGVNDIFINVRTH-RMRDSGLFSGVDTEAERIGRTVGKAGPAVLSSLS 737
QY 754 EAICPFLGALTMPAVRTFALTSGLAVILD FLOWMSAFVALLSDSKRQZASRLDVCVVCCV 813
Db 738 ESACCLIGALSMPAVNTFALPAATALAINFLVQITFTFVALMTLDSNRVLRARLDVACCI 797

QY 1040 SD-----GOVLASFMAHKLKNSQDYTEALRAARELANITADLRKVFPTDPAFVFPY 1095
DB 975 IDVGMSTVQDSYFMOYSTTSTTSEBFYSQLRVRRISGEINAMFKE---NNVDABIFAY 1031
QY 1096 TITNVFEQVLTLLPGLFMLSCLVPTFAVSCLLGLDLRSLGLNLLSVMLIVTVGF 1155
DB 1032 CVFYIYEQVLTWGDAMFSLGSLVAIFLVLTLLITGLDITSTFVLFVVICILNMLGM 1091
QY 1156 MALWDISYNAVLINLVSAGMSVFEVSHITRSFALSTKPTMLERAKEATISMGSAVPAG 1215
DB 1092 MWAMSINLNAISLVNLVVCVIGVEFVAHVRSFK-RAEGTAQERARHSINVTGSSVLSG 1150
QY 1216 VAMTNLPGLVLGLAKAQIQLPFFRLNLTLLGLLHGLVFLVILSVYGVDPNPAAL 1275
DB 1151 IITLTFAGIWLGLFSNSQIFQVFFRMYLGIIVLGAAGHLIILPVLLSLGPPQK----L 1206
QY 1276 EQKRAEEAANVMVAS 1291
DB 1207 ARSSGAETASITITT 1222

RESULT 98

US-09-614-150A-2679
; Sequence 2679, Application US/09614150A
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150A
; PRIOR FILING DATE: 2000-07-11
; PRIOR FILING DATE: 2000-07-11
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2679
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150A-2679

Query Match 27.0%; Score 1864.5; DB 26; Length 1223;
Best Local Similarity 33.5%; Pred. No. 7.2e-154;
Matches 428; Conservative 237; Mismatches 498; Indels 113; Gaps 27;
QY 62 PARKITG-DHLLLOKICPRLY-----TGNP-TQACCSAKOLVSLASLSITKALLTRC 113
DB 14 PARLNSPTSEAFKRCPLMKYKESGSEDELSLCCDAQIETWESGLSQADGVFSRC 73
QY 114 PACSDNFVNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQS----Y 169
DB 74 PTCRNMALTVCACTAKNHTLFLAYNDTN-----DAGVDYVKYIDYRLTDDTVSKIY 127
QY 170 DCSRSRVPAANTLAVGTWCGVYGSALCNAQRWLNFGQDTGNGLAPLDTITFHLBPGQAV 229
DB 128 NSCIGTQHTQTRPAMDLCGCSYNAKTCNRRRWYBFMGDVGSDYVFPQINRYKWSDEABE- 186

QY 230 GSGIOPLANEGBVARCNESQDDVATCSODCAASCFAIARPOALDSTFYLGQMEG--SLVL 287
DB 187 GSENIYLDLSPLKCGESYEDSYA-CACIDCBESCLPTDAPTGPDELWKIAGLVGVIFILA 245
QY 288 IILICSVFAVTVILLVGFVAVPARDKSKWDPKGTSLSDKLSFSTHTLLGQF-----FQ 343
DB 246 LTIACALSPFIIPWAGFKTSAP-----SVCMTPLTFGEFFYHGFR 285
QY 344 WGTWVASWPLTLLVLSVIPPVALAAGLVFELTDTDPVELWSAPNSOARESEKAFHDQHFGP 403
DB 286 WGTFCAPKHPVIVLALCSWAIAGLSFGIRTWITTDPELVAGEESQRIEKDIFDQHFGP 345
QY 404 FPRITQVILTAPERRSYRYSLLGPKNFSGLDLDLLELELELELELELELELELELELELE 463
DB 346 FYRTNQMFVAVNTQVTFTHET-SNGVLNFGPAPEYFNFLAKVFELODSIMKLG-AD 399
QY 464 NISLQDICYAP-LNPDNSTLYD-CCINSLLQVFNQNRITLLLTANQTLGQTSQVNDKOH 521
DB 400 NEGLDKICYAPVLMAGETPTVDRCAIQSVYGYFQHDMDRF---ENSVDNSNNYITINLQ 456
QY 522 FLYCANAPLTFKDGITALALSMDADYCAPVPPFLAIGGY-----KGKDYSEAEALIMTFS 575
DB 457 LEDCLRVPM-MED-----CFGTFGGPIEPGIAVGMKPVAVGEDPDMLATGLVITFL 508
QY 576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFQORMMAGMFQVTTAERSLEDEINRTTARDLPI 635
DB 509 GRNY-NDESKLRPNKWEKLFVDLFDYK---SDRLDIAYMAERSIQDAIVELSEGEVST 564
QY 636 PATSVIVIPLYTSLALGSYSWSRVWDSKATLGLGGVAVVILGAVMAANGPFSYLGIRSS 695
DB 565 VVISYVWVYVAIALGHIRSCRGFLRESRIMLAIGGIVIVLASVVCISLGFYGLDVTVT 624
QY 696 LVILQVPELVLSVGDADNIFIVLEVORLPRPGEPRVHIGRALGRVAPSMCLLSSEA 755
DB 625 MIAIEVLPVLAVGVDNIFIMVHTYQRLDHSKFTTHEAIGATQVGSFSLQTAGSEM 684
QY 756 ICFFLGALTMPAVRTFALTSLGLAVLDFLQMSAFVALLSDSKRQESRLDVCCKVK- 814
DB 685 ACFAIGCISDMPAVKTFAVYAAIILDFLQITAFVALMAIDEXYLDGLRDLMLCCVKS 744
QY 815 -----PQELPPPGQEGILLGFPKAYAPFLHWTIRGVVLLLFLALFGV 859
DB 745 GGGKINDEGDGVDPRPKEV-----GLLETFLKPNFYSFPLLSKPKVSVLITFTVITCL 797
QY 860 SLYSMCHISVGLDDQELALPKDSYLDYFLFLNRYFEVGPVYFVTLLGVNFSSEAGMNAI 919
DB 798 SLWVTPSIEKGLDQEMSPKNSHVVKYFRYMDLLAMGAPVYVWLKPGVLYSEPLQONLI 857
QY 920 CSSAGCNNSFTQKIQYATEPPEQSYLAIPASSWDDDFIDMLTPSSCCRLYISGPNKDF 979
DB 858 CGGVECNNSLSVQLYTQAYPEITSLARPASSWDDYIDMLAISDCCKYNT---TCGF 914
QY 980 CPSTVNSLNLKNCMSITMGSVRSPVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVNLT 1039
DB 915 CSSNSKSEDCPCERGFTEGLRPAETFNKYIPYFLPDLPAECAKAGRASVADAVIT 974
QY 1040 SD-----GOVLASFMAHKLKNSQDYTEALRAARELANITADLRKVFPTDPAFVFPY 1095
DB 975 IDVGMSTVQDSYFMOYSTTSTTSEBFYSQLRVRRISGEINAMFKE---NNVDABIFAY 1031
QY 1096 TITNVFEQVLTLLPGLFMLSCLVPTFAVSCLLGLDLRSLGLNLLSVMLIVTVTVGF 1155
DB 1032 CVFYIYEQVLTWGDAMFSLGSLVAIFLVLTLLITGLDITSTFVLFVVICILNMLGM 1091
QY 1156 MALWDISYNAVLINLVSAGMSVFEVSHITRSFALSTKPTMLERAKEATISMGSAVPAG 1215
DB 1092 MWAMSINLNAISLVNLVVCVIGVEFVAHVRSFK-RAEGTAQERARHSINVTGSSVLSG 1150
QY 1216 VAMTNLPGLVLGLAKAQIQLPFFRLNLTLLGLLHGLVFLVILSVYGVDPNPAAL 1275
DB 1151 IITLTFAGIWLGLFSNSQIFQVFFRMYLGIIVLGAAGHLIILPVLLSLGPPQK----L 1206
QY 1276 EQKRAEEAANVMVAS 1291

; SEQ ID NO 514
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Drosophila
US-60-161-932-514

Query Match 27.08; Score 1864.5; DB 45; Length 1223;
Best Local Similarity 33.58; Pred. No. 7.2e-154;
Matches 428; Conservative 237; Mismatches 498; Indels 113; Gaps 27;
QY 62 PARKITG-DHLIILQIKCPRLY-----TQPN-TQACCSAKQLVSLRSLITKALLTRC 113
DB 14 PARPLNSPTSEAFKCPMLYKEYKGESEDELSLCCDAQAQIETWESGLSQDGFPSRC 73
QY 114 PACSDNFVNHLCHNTCPNOSLFINVTRVLAQAGQLPAVVAYEAFYQHSFABQS-----Y 169
DB 74 PTCRNNALVCAVCAKNTLFLTAINDTN-----DAGVDYVVKYIDYRLTDDTVSKIY 127
QY 170 DSCSRVRVPAATLAVGTGCGVGSALCNAQRMLNFGDGTGNGLAPLDDITPHLLEPGQAV 229
DB 128 NSCIGIQTGTGRPAMDLCGCSYNAKTCNRRRYEFMGDVSGDYVPFQINYNKWSDEABE- 186
QY 230 GSGIQLNEGVARCNESQGDVATCSQDCAACAPALARPQALDSTFYLGQMPG--SLVL 287
DB 187 GSNEIYLDLSPKCGESYDSA-CACIDCEESCLPTDAPTGDELWKIAGLIGVTFILA 245
QY 288 IILCSVPFVAVTILLVGRFVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQF----FOG 343
DB 246 LIITACALSPFIFWGAFGKTSAP-----SVCMPPTLFGBFFYHGPR 285
QY 344 WGTWASWPLTILVLSVIPVALLAGLVFTELTDPVELWSAPNSQARSEKAFHDQHPG 403
DB 286 WGTFCARHPVILALCSWAIAGLSFGIRYMTITDPVELWAGESQTRIEKDYFDQHPG 345
QY 404 FFRTNQVILTAPNRSSRYDLSLGPKNFSGIILDLLELLELQERLRLHLOVMSPEAQR 463
DB 346 FYRTNQMFYKAVNQTYFTHET-SNGVINFGPAPFAYNPLKEVFELOQSIMKLG-AD 399
QY 464 NISLQDICVAP-INPDNTSLYD--CCINSLLQYQNNRTLLLLPANQTLMGQTSQVDWKDH 521
DB 400 NEGLDKICVAPVLMAGETPTVDRCAIQSVVGYFQHDMDRF---ENSVDSSNNYINYLQ 456
QY 522 FLYCANAPLTFKDTALALSCMADYCAPVFPFLAIGY-----KGDYSEABALIMTFS 575
DB 457 LEDCLRVPM-MED-----CFGTGPGIEPGIAGVGMKPAVGEDPDYMLATGLVLTEL 508
QY 576 LNNYPAGDPRLAQAKLWEBAFLTEEMRAFORRMAGMFQVTFATERSLEDEINRTAEDLPI 635
DB 509 GRNY-NDESKLEPNMKWEKLFVDFLDYK---SDRLDIAYMAERSIQDAIVELSEGEVST 564
QY 636 FATSIVIFLYISLALGSSYSSMSRVNVDKATILGLGGAVALGAVNMAAMGFFSYLGIRSS 695
DB 565 VVISYVVMFVVAIALGHIRSCRGPLRESRIMLAIGGIVIVLASVVCSLGFWCYLDVTTT 624
QY 696 LVILQVVPFLVLSVGADNIFIFVLEFQRLPRPGEPREVHIGRALGRVAPSMMLCSLSEA 755
DB 625 MIAEIVIPFLVLAAGVDNIFIMVHTYQRLDHSKFKTTHEAIGEAIGQVGFSLIQTAGSEM 684
QY 756 ICFPLGALTMPMPARTFALTGLAVILDFLQMSAFVALLSLDSKQESRLDVCCKVK- 814
DB 685 ACFAIGCISDMPAVKTFAMTAAILLDLFLQITAFVMAIDERRYLDRGLDMLCCVKS 744
QY 815 -----POELPPGQGGEGLLGPFQKAYAPFLHWHITRGVVLFLFLALFGV 859
DB 745 GGGKINDEGDGVDREPKV-----GLLETLPKNFYSPLLSKPKVSVLLIFTVITCL 797
QY 860 SLYSMCHISVGLDQELALPKDSVLLDYFLFLNRYFVGVAPVYFVTTLVGNFSSRAGMNAI 919
DB 798 SLMTVPSIEKGLDQEMSPKNSHVVKYFRYMYDALLANGAPVYVVLKPLGNYSRPLQONLI 857
QY 920 CSSAGCNNEFSFTQKIQAYTEFPQSYLAIPASSWDDDFIDWLTPSSCCRLYISGPNKDXF 979
DB 858 CGGVECNNSLSVQLYTOAQYPEITSLARPASSWLDYIDWLAISDCCCKYNVT---TGGF 914

Search completed: April 11, 2006, 00:56:31
Job time : 594 secs

QY 980 CPSTVNSLNLCKNCMSITMGSVRPSVEQFHKYLFWFLNDRPNIKCPKGGKLAAYSTSVNLT 1039
DB 915 CSSNSKSEDCLPCERGFTENGRLPDAETFNKYIPYFLFDLPDAECACAGRASYADAVIYT 974
QY 1040 SD-----GOVLASRPMAYHKPLKNSQDYTEALRAARELANITADLRKRVPGTDPAFVFPY 1095
DB 975 IDDVGMSTVQDSYFMQYSTTSTTSEEFYSQLREVRRISGEINAMFKE---NNVDABIFAY 1031
QY 1096 TITNVFVEOYLTILPGLFMLSCLVPTPAVSCILLGLDLRSGLNLISIVMLIVDTVGF 1155
DB 1032 CVFYIYEQYLTITWGDAMFSLGMSLVAIFLVLTILITGLDITSTFIVLFWVICILINMLG 1091
QY 1156 MALWDISYNVSLINLVSVMGSEVFSVSHITRSFAISTKPTWLERAKEATISMGSAVFAG 1215
DB 1092 MWAMSLNLAISLVNLVVCVIGVGFVAHVRSFK-RAECTAQERARHSLNVTGSSVLSG 1150
QY 1216 VAMTNLPGLVLGLAKAQLIOIPFRLNLLITLLGLLHGLVFLPVLISYVGPDPNPAAL 1275
DB 1151 ITLTKEPAGIVLWGFNSQIQVFFRMYLGIIVLIGAAGHLILPVLISLILGPPQK----L 1206
QY 1276 EOKRAEEAVALVMVAS 1291
DB 1207 ARSSGAEPASITITT 1222

THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:47:36 ; Search time 78 Seconds
(without alignments)
2379.065 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGATSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 592492 seqs, 139314628 residues

Total number of hits satisfying chosen parameters: 592492

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Pending Patents AA New:*

- 1: /SIDSS5/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /SIDSS5/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /SIDSS5/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /SIDSS5/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /SIDSS5/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /SIDSS5/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /SIDSS5/ptodata/2/paa/US11_NEW_COMB.pep.*
- 8: /SIDSS5/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6909	100.0	1332	6	US-11-301-094-6
2	6872.5	99.5	1359	6	US-11-270-796-22
3	5407	78.3	1333	6	US-11-270-796-3
4	2402.5	34.8	1278	1	PCT-US06-05584-822
5	2402.5	34.8	1278	6	US-11-191-274A-129
6	2402.5	34.8	1278	6	US-11-191-274A-130
7	2402.5	34.8	1278	7	US-11-385-692-2400
8	2402.5	34.8	1278	7	US-11-385-692-2401
9	1341.5	19.4	1383	6	US-11-301-094-2
10	1046	15.1	1274	6	US-11-301-094-4
11	782.5	11.3	419	6	US-10-953-349-21065
12	588.5	8.5	967	6	US-10-461-673-16739
13	583.5	8.4	1182	6	US-11-332-764-2
14	583.5	8.4	1182	6	US-11-337-244-149
15	482	7.0	891	8	US-60-772-265-1197
16	455	6.6	1358	8	US-60-772-265-241
17	437	6.3	831	7	US-11-360-355-133805
18	349	5.1	783	6	US-10-461-673-16750
19	325	4.7	204	7	US-11-360-355-141351
20	281	3.8	575	7	US-11-360-355-120785
21	252	3.6	505	6	US-10-461-673-16871
22	229.5	3.3	465	7	US-11-360-355-141367
23	225	3.3	542	6	US-11-214-063A-1670
24	222.5	3.2	422	7	US-11-360-355-141365
25	219	3.2	1137	6	US-10-461-673-16889

Sequence 10541, A	1330	3.2	219	6	US-10-461-673-10541
Sequence 141352, A	126	2.8	190	7	US-11-360-355-141352
Sequence 3548, Ap	632	2.2	152.5	7	US-11-293-697-3548
Sequence 127069, A	164	2.0	141	7	US-11-360-355-127069
Sequence 120169, A	484	2.0	139.5	7	US-11-360-355-120169
Sequence 140206, A	252	2.0	135.5	7	US-11-360-355-140206
Sequence 7050, Ap	808	1.8	126	6	US-10-724-972B-7050
Sequence 149589, A	171	1.8	123.5	7	US-11-360-355-149589
Sequence 20, Appl	1043	1.7	120.5	6	US-10-536-606-20
Sequence 2044, Ap	768	1.7	118.5	6	US-11-203-806A-11
Sequence 1692, Ap	788	1.7	117.5	6	US-11-214-063A-2044
Sequence 1828, Ap	619	1.7	117.5	8	US-60-732-162-1828
Sequence 46, Appl	619	1.7	116.5	6	US-11-312-958-46
Sequence 12167, A	619	1.7	116.5	6	US-10-461-673-12167
Sequence 45206, A	488	1.7	114	8	US-60-752-355-45206
Sequence 2874, Ap	985	1.7	114	6	US-11-293-697-2874
Sequence 152528, A	697	1.6	112.5	6	US-10-703-799B-226
Sequence 21, Appl	958	1.6	111	8	US-60-781-953-21
Sequence 2027, Ap	494	1.6	110.5	6	US-11-045-004-2027
Sequence 2398, Ap	451	1.6	109.5	6	US-11-045-004-2398
Sequence 73134, A	758	1.6	109.5	7	US-11-056-355B-73134
Sequence 73132, A	827	1.6	109.5	7	US-11-056-355B-73132
Sequence 33716, A	3979	1.6	109.5	8	US-60-752-355-33716
Sequence 39891, A	474	1.6	108	6	US-60-752-355-39891
Sequence 5410, Ap	901	1.5	107	6	US-10-724-972B-5410
Sequence 1237, Ap	629	1.5	106.5	6	US-10-567-867-1237
Sequence 16314, A	435	1.5	106	8	US-60-752-355-16314
Sequence 3216, Ap	924	1.5	106	7	US-11-027-399A-3216
Sequence 3216, Ap	924	1.5	106	7	US-11-028-099A-3216
Sequence 3216, Ap	924	1.5	106	7	US-11-027-878A-3216
Sequence 3216, Ap	924	1.5	106	7	US-11-028-291A-3216
Sequence 15654, A	637	1.5	105.5	6	US-10-917-503B-15654
Sequence 272, App	783	1.5	105.5	8	US-10-703-799B-272
Sequence 2802, Ap	1077	1.5	105	8	US-60-751-420-2802
Sequence 119639, A	308	1.5	105	7	US-11-360-355-119639
Sequence 69789, A	1080	1.5	105	7	US-11-056-355B-69789
Sequence 69788, A	1097	1.5	105	7	US-11-056-355B-69788
Sequence 129465, A	1128	1.5	104	7	US-11-056-355B-129465
Sequence 149551, A	238	1.5	104	6	US-11-360-355-149551
Sequence 4902, Ap	510	1.5	104	6	US-10-724-972B-4902
Sequence 191, App	692	1.5	104	6	US-10-159-257C-191
Sequence 69565, A	757	1.5	104	7	US-11-056-355B-69565
Sequence 69564, A	771	1.5	104	7	US-11-056-355B-69564
Sequence 8918, Ap	826	1.5	103.5	7	US-11-056-355B-8918
Sequence 7153, Ap	618	1.5	103.5	1	PCT-US06-00964-7153
Sequence 7153, Ap	618	1.5	103.5	6	US-11-330-403-7153
Sequence 25099, A	759	1.5	103.5	8	US-60-752-355-25099
Sequence 4330, Ap	1136	1.5	103	6	US-11-045-004-1583
Sequence 246, App	463	1.5	103	8	US-60-752-355-4330
Sequence 37989, A	470	1.5	102.5	6	US-10-703-799B-246
Sequence 1806, Ap	472	1.5	102.5	6	US-60-752-355-37989
Sequence 208, App	538	1.5	102.5	6	US-10-498-451-1806
Sequence 25298, A	752	1.5	102.5	8	US-60-742-219-208
Sequence 14490, A	304	1.5	102	1	PCT-US06-07642-14490
Sequence 14490, A	304	1.5	102	8	US-60-658-984A-14490
Sequence 2241, A	569	1.5	102	8	US-60-752-355-2241
Sequence 1750, Appl	1364	1.5	102	6	US-10-498-451-1750
Sequence 41, Appl	634	1.5	101.5	6	US-60-781-306-41
Sequence 12928, A	397	1.5	101	6	US-10-917-503B-12928
Sequence 2249, Ap	397	1.5	101	6	US-10-533-519-2249
Sequence 260, App	475	1.5	101	6	US-10-206-921A-260
Sequence 260, App	475	1.5	101	6	US-10-184-614A-260
Sequence 1676, Ap	475	1.5	101	8	US-60-732-162-1676
Sequence 7829, Ap	556	1.5	101	8	US-60-752-355-7829
Sequence 3734, Ap	607	1.5	101	7	US-11-174-307B-3734
Sequence 19, Appl	1016	1.5	101	1	PCT-US05-46487-19

99 101 1.5 1016 6 US-11-314-892-19 Sequence 19, Appl
100 101 1.5 1016 7 US-11-375-551-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-11-301-094-6
; Sequence 6, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-094-6

Query Match 100.0%; Score 6909; DB 6; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDLHLLLOKICRPLTYGPNTOACCSAKQLVSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDLHLLLOKICRPLTYGPNTOACCSAKQLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSYDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSYDSCSRVRPAA 180
QY 181 ATLAVGTCGVGSAALCNQRLNFGDGTGNGLAPLDITPHLLEPQAVGSGIQLNEGV 240
DB 181 ATLAVGTCGVGSAALCNQRLNFGDGTGNGLAPLDITPHLLEPQAVGSGIQLNEGV 240
QY 241 ARCNESSQDDVATCSQDCAASCAPAIARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESSQDDVATCSQDCAASCAPAIARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVPAIPARDKSKVMDPKGTSLSDKLSFSFTHLLQFPQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVPAIPARDKSKVMDPKGTSLSDKLSFSFTHLLQFPQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFELTTPVELMSAPNSQARSEKAFHDFGPFPTNQTIVLTAPNRSY 420
DB 361 IPVVALAAGLVFELTTPVELMSAPNSQARSEKAFHDFGPFPTNQTIVLTAPNRSY 420
QY 421 RYDSSLGPKNFGSGLDLDLLELLELQRLHLQVWSPQARNISLQICVAPLNPDNT 480
DB 421 RYDSSLGPKNFGSGLDLDLLELLELQRLHLQVWSPQARNISLQICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGTALAL 540
QY 541 SCHADYGAPVFPPLATGGYKDKDYSEAEALIMFTFSNNYPAGDPRLAQAKLWEEAELEM 600
DB 541 SCHADYGAPVFPPLATGGYKDKDYSEAEALIMFTFSNNYPAGDPRLAQAKLWEEAELEM 600
QY 601 RAFORMMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660

DB 601 RAFORMMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATILGLGGVAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFLVLE 720
DB 661 MYDSKATILGLGGVAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFLVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSISEAICPFLGALTTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSISEAICPFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPPQGSGLLLGFPQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPPPQGSGLLLGFPQKAYAPFL 840
QY 841 LHMITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLPLNRYFEVGAPV 900
DB 841 LHMITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLPLNRYFEVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNSFTQIKIYATEFFPQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNSFTQIKIYATEFFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFPCSTVNSLNCNCSITMGSVRPSVEQFHKYLPMFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFPCSTVNSLNCNCSITMGSVRPSVEQFHKYLPMFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQVILASRPMAYHKPKNSODYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTVNLTSQVILASRPMAYHKPKNSODYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEVFPYTTINVFYEQVLTTLPEGLFMLSCLVPTFAVSCILLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEVFPYTTINVFYEQVLTTLPEGLFMLSCLVPTFAVSCILLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVPGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSIVMLVDTVPGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGIILVLGLAKAQLIOIFPFRNLNLTLLGLLHGLVFLPV 1260
DB 1201 AKEATISGMSAVFAGVAMTNLPGIILVLGLAKAQLIOIFPFRNLNLTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
DB 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 2

US-11-270-796-22
; Sequence 22, Application US/1120796
; GENERAL INFORMATION:
; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; CURRENT FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-270-796-22

Query Match 99.5%; Score 6872.5; DB 6; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

```
QY 1 MAEAGLRGWLWALLRLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASISITKALLTRCPACSDNP 120
Db 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASISITKALLTRCPACSDNP 120
QY 121 VNLHCHNTCPNQSLEFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAA 180
Db 121 VNLHCHNTCPNQSLEFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAA 180
QY 181 ATLAVTGTCGVYGSALCNAQRMWLNFGQDGTGNGLAPLDTIFHLLPEQAVGSGIQPLNEGV 240
Db 181 ATLAVTGTCGVYGSALCNAQRMWLNFGQDGTGNGLAPLDTIFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVATCSQDCODCAASCPAIPARQALDSTFYLQMGPSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCODCAASCPAIPARQALDSTFYLQMGPSLVLIILCSVFAVVTI 300
QY 301 LLVGPVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLVS 360
Db 301 LLVGPVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLVS 360
QY 361 IPVALLAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSY 420
Db 361 IPVALLAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSY 420
QY 421 RYDSLILGPNFSGIILDLLELLELRLHRLQWSPEAQNSLQDICYAPLNPDNT 480
Db 421 RYDSLILGPNFSGIILDLLELLELRLHRLQWSPEAQNSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKQHFYLCANAPLTKDGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKQHFYLCANAPLTKDGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKDYSEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKDYSEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFORMMAGMFWTFAERSLEDEINRTTAEDLPFATSVIIVIFLYISIALGYSWSRV 660
Db 601 RAFORMMAGMFWTFAERSLEDEINRTTAEDLPFATSVIIVIFLYISIALGYSWSRV 660
QY 661 MYDSKATLGLGVAVVILGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVILGAVMAAMGFYSYLGIRSSLVILQVVPFLVLSVADNIFIFVLE 720
QY 721 YORLPRRPGEPREHVHIGRALGRVAPSMILCSLSEATCFPLGALTMPAVRTFALTSGLAV 780
Db 721 YORLPRRPGEPREHVHIGRALGRVAPSMILCSLSEATCFPLGALTMPAVRTFALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSDSKRQEASRLDVCCKVQELPPQEGGLLIGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSDSKRQEASRLDVCCKVQELPPQEGGLLIGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLYGNFSSEAGNNAICSSAGCNPFSTQKIQYATEFPQSYLEIAIPASSWVDDFIDW 960
Db 901 YFVTTLYGNFSSEAGNNAICSSAGCNPFSTQKIQYATEFPQSYLEIAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCNKCMSITMGSVRPSVBOFHKLIPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCNKCMSITMGSVRPSVBOFHKLIPWFLNDRP 1020
QY 1021 NIKCPKGGIAAYSTSVNLTSDGQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGGIAAYSTSVNLTSDGQVLDTVAILSPRLSEYSGTISAHCNLYLLDSASRFMAYH 1080
```

```
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYTTITNVFEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYTTITNVFEQYLTILPEGL 1140
QY 1114 FMLSICLVPTFAVSCLLIGLRLSGLNLLSVMLIVDTVGMALWDISYNAVSLNLVS 1173
Db 1141 FMLSICLVPTFAVSCLLIGLRLSGLNLLSVMLIVDTVGMALWDISYNAVSLNLVS 1200
QY 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAFVAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAFVAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIPFFRLNLLITLLGLLHGLVFLPVILSYVGPDPVPALAEQKAEBAVAAMVASCP 1293
Db 1261 LIQIPFFRLNLLITLLGLLHGLVFLPVILSYVGPDPVPALAEQKAEBAVAAMVASCP 1320
QY 1294 NHPSRVSTADNIVNHSPEGSIKGAGATSNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIVNHSPEGSIKGAGATSNFLPNNGRQF 1359

RESULT 3
US-11-270-796-3
; Sequence 3, Application US/11270796
; GENERAL INFORMATION:
; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; TITLE OF INVENTION: RESTORATION OR MIMICRY PP p16 Ink4a ACTIVITY
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; CURRENT FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 3
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-270-796-3
```

```
Query Match 78.3%; Score 5407; DB 6; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGWLLWALLRLRLAQSEPYTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA 63
Db 3 AAWQGWLLWALLRLNSAQGELYTPTHKAGCTFYBCKGNPELSGLTSLNSICLSNTPA 62
QY 64 RKITGDHLLILLOKICPRLYTGPN-TOACCSAKQLVSLASISITKALLTRCPACSDNFVN 122
Db 63 RHVTGDHLLALQVCPRLYNGENDTYACCSYQLVSLDSSLSITKALLTRCPACSDNFVS 122
QY 123 LHCHNTCPNQSLEFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAAAT 182
Db 123 IHCHNTCPDQSLFINVTVVQDPGOLPAVVAEAFYQHSFABQESCSRVRIPAAAS 182
QY 183 LAVGTGTCGVYGSALCNAQRMWLNFGQDGTGNGLAPLDTIFHLLPEQAVGSGIQPLNEGVAR 242
Db 183 LAVGSMCGVYGSALCNAQRMWLNFGQDGTGNGLAPLDTIFHLLPEQALADMKPLDKITP 242
QY 243 CHNESQDDVATCSQDCODCAASCPAIPARQALDSTFYLQMGPSLVLIILCSVPAVVITLL 302
Db 243 CHNESQGBSAAACSCQDCAAACSVIIPPPALRPSFYNGRMPGWLIIITFAVFLLSVVL 302
QY 303 VGRFVAPARDKSKMVDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSPV 362
Db 303 VYLRVASNRKNTAGSQEAPNLPRRRPSPHTVLRFPFESWGTVRVASWPLTILVLSPIV 362
QY 363 VVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSYRY 422
Db 363 VTALSGLTFFIELTTDPPVELWSAPKQARKEKAFHDEHFGPFFRTNQIFVTAKNRSSYKY 422
```

```
QY 423 D5LLGPKNFSGILDLLLELLELOERLRHLQVMSPEAQRNISLODICYAPLNPNTSL 482
Db 423 D5LLGPKNFSGILSDLLLELLELOERLRHLQVMSHEAQRNISLODICYAPLNPHTSL 482
QY 483 YDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTLALASC 542
Db 483 TDCCVNSLLQYFQNNHTLLLTANQTLNGQTSVDWKDHFLYCANAPLTYKDGTLALASC 542
QY 543 MADYGAPVPPFLAIGYKGKDYSEABALIMTFSLNYPAGDPRLAQAKLWBEAFLEEMRA 602
Db 543 IADYGAPVPPFLAVGGYQGTSEABALITFSINNYPADDPMAHAKLWBEAFLEEMQS 602
QY 603 FORRMAGHFOVTFEATERSLEDEINTEADLPIFATSVIVFLYISLALGSYSSSRVUV 662
Db 603 FORSTADKFOIAFSAERSLEDEINRTIQDLVFAISLIVFLYISLALGSYSSSRVAV 662
QY 663 DSKATILGCGVAVLGVAMMAAGFFSYIGRSVLIVQVPLVSLVAGADNIFIFVLBYQ 722
Db 663 DSKATILGCGVAVLGVAMVANGFFSYLGVPSLVIQVPLVAVAGADNIFIFVLBYQ 722
QY 723 RLPRRPGPREVHI GRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMPGEOREAHIGRTLGSVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSGLAIF 782
QY 783 DELLQMSAFVALLSLSKQREASRLDCCVXPQELPPQCGEGLLGGFFQKAYAPFLH 842
Db 783 DELLQMTAFVALLSLSKQREASRPDVCFSRNLPPPKQEGELLGFFRKIYTFPFLH 842
QY 843 WITRGVLLFLFALGVSILYSCHISVGLDQDLALPKDSYLLDYFLFLNRYEVEGAPVYF 902
Db 843 RIRPVLVLLFLVFGANILYMCNLSVGLDQDLALPKDSYLLDYFLFLNRYLEVGPVYF 902
QY 903 VTLTGVNFSSEAGMAICSSAGCNPFSTQKIQTATEPPEQSYLAIPASSWVDDPIDMLT 962
Db 903 DTTSGYNFSTEAGMAICSSAGCESFSLTKIQTAYASEFPNQSYVAIAASSWVDDPIDMLT 962
QY 963 P-SSCCRLYISGPNKDKCPTVNSLNCNKMSITMGSVRPSVQPHKYLPMFLNDRPN 1021
Db 963 PSSCCRIYTRGPHKDEPCPTDTDFNCLKNCNRTLGPVRPTTQFHKYLPMFLNDRPN 1022
QY 1022 IKCPKGLAAYTSVNLTSDDQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADIR 1081
Db 1023 IRCPKGLAAYTSVNLSSDGOI IASQFWAYHKLKNSQDFTEALRASLLAANITAEIR 1082
QY 1082 KVPGTDPAPVPPYITNVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLN 1141
Db 1083 KVPGTDPNPEVPPYITSNFYQYLTLPPEGITLALCFVPTFVVCYLLGLDLRSGLN 1142
QY 1142 LLSIVMILVDTYGFALWDISYNAVSLINLAVAGMSVEFVSHITRSPAISTKPTMLERA 1201
Db 1143 LLSIIMILVDTITGLMAVWGISYNAVSLINLAVAGMSVEFVSHITRSPAISTKPTMLERA 1202
QY 1202 KEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIFFPRMLNLTLLGLHGLVFLPYI 1261
Db 1203 KDATIFMGSAVPAVAMTNFPGILVLGLAQAQLIQIFFPRMLNLTLLGLHGLVFLPV 1262
QY 1262 LSYVGDVNPALALEQKRAEEAAVAVMVASCPNHPSSRVSTANIIYNHSHFEGS - IKGGA 1320
Db 1263 LSYLGDVNPQALVLEBKATEA - AMVSEPCQYPPPADANTSDYVNGFNPFEIPEINA 1321
QY 1321 ISNFLPNNGROF 1332
Db 1322 ASSSLPKSDQKF 1333
```

RESULT 4

```
PCT-US06-05584-822
; Sequence 822, Application PC/TUS0605584
; GENERAL INFORMATION:
; APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION,
; TITLE OF INVENTION: ASSESSMENT, PREVENTION, AND THERAPY OF CANCER
```

```
; FILE REFERENCE: DFS-064.25(25992-6425)
; CURRENT APPLICATION NUMBER: PCT/US06/05584
; CURRENT FILING DATE: 2006-03-02
; PRIOR APPLICATION NUMBER: 60/690,064
; PRIOR FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: 60/654,227
; PRIOR FILING DATE: 2005-02-17
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 822
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US06-05584-822
```

```
Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGMILLMALLRLAQSEPPYTTIHQPCYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
Db 4 RGLALGLLLLLLCPAQFSQ-----SCVWYGECCGIAYGDKRYNCBYSG-----46
QY 60 NTPARKITGDRHLILLQKICPRLYTGPNTOACCSAKQLVSLBASISITKALLTRCPACSDN 119
Db 47 --PKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDNLQLPQLFLSRCFCSCFN 103
QY 120 FVNHLCHNTCPNOSLFINVTR----VAQLGAGQLPAVAYEAFYQHSHFASQSYDSCSR 175
Db 104 LNLFCELTCSPROQFLNVTATEDYDVPVNTQNTKYNKELQYVYGVSFANAMYNACRDV 163
QY 176 RVPAATLAVCTMCGVYGSALCNAORWLNFGDGTGGLAPLDIT-----PHLLEPGA 228
Db 164 EAPSNDKALGLLCKODADA-CNATNMEYMFNKGNGOAPPTITPVFSDPPVH-----215
QY 229 VSGIQLPNEGVARCNESQDDVATCSQDCAASCFAIARPO-----ALDSTFYLG 279
Db 216 ---GWEPNNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSVFAVTILL-----VGFVRVAPARDKSWDPKKGTSLS 326
Db 273 WITYMAFLVFPFAPFAVWCYRKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQRFQGWGTWVASWPLTILVLSPVVALAAGLVFETLTTPDVELWSAP 386
Db 325 DPVSAAPFGCLRLRLETRWGSFCVRNPGCVIFESLVFITACSSGLVFRVTTNPDLWSAP 384
QY 387 NSQARSEKAFDQHFPPFRTNQVILTAPNRSYRYDSLLGLPK-NFGSILDLDLLELL 445
Db 385 SSQARLEKEYFDQHFPPFRTTEQLIIRAPLTDKHIYQVPVSGADVFPGLDIQLHQVL 444
QY 446 ELQERLRHLQVMSPEAQRNISLODICYAPLNPNTSLVDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVLNFQNSHSLDHKK 499
QY 506 NOTLMGQTSQVDWKDHFLYCANAPLTFKDGTLALASCMAHYCAPVFPFLAIGYKGKYS 565
Db 500 GDDFF---VYADITHFLYCVRAPASLNDTSLLDHPCLTGTFGCPVFPVLVGGYDDQNTN 556
QY 566 EAEALIMTFSLNYPAGDPRLAQAKLWBEAFLEEMRAFQRRMAGMFQVTFEATERSLEDBI 625
Db 557 NATALVITFPVNNYNDTEKLQRAQWEKEFINFVKYKN---PNLTISFTAERSIETDEL 613
QY 626 NRTTAEDLPFATSVIVFLYISLALGSYSSSRVWVDSKATILGCGVAVLGVAMMAAG 685
Db 614 NRESDDVFTVVISYAIMFLYISLALGHKSCRRLLVDSKSLGSLAGILIVLSSVACSJG 673
QY 686 PFSYLGIRSSVLQVPPFLVSLVAGADNIFIFVLBYQRLPRRPGPREVHI GRALGRVAP 745
Db 674 VPSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQORDERLQGETLDOQLGRVLGEVAP 733
QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSGLAIVLDFLLQMSAFVALLSLSKQREAS 805
```

```
Db 734 SMFLSFSFSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLLOITCFVSLLLGLDIIKQKRN 793
QY 806 RLDVCCVKPQELPPGQ-GEGLLLGFFOKAYAPFLLHWITRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESLCFRFFKNSYSPLLKKDWMRPVIAIFVGLSPSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLPLNRYFEVGAUVYFVTLGYNFSSRAGMAICSSAG 924
Db 854 NKVDIGLDQSLSPDDSYMDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVGCGMG 913
QY 925 CNMFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDWLTP--SSCCRLYISGPNKDFPCPST 983
Db 914 CNMDSLVQOIFNAAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNMCSTT-MGSVRPSVEQFHKYLPWFILNDRPNIKCPKGGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR-CRPLTPBEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSSAVNLLGH 1029
QY 1041 DGOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMTVHTVLTQTSADFIDALKARLIASNTV-ETMGINGS--AIRVFPYSVFTV 1086
QY 1101 FYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLSRLNLLSIVMLVDTVGFMAWD 1160
Db 1087 FYEQYLTIIDDTIFNLGVSGLAIFLVTWLLGCELWSAVIMCATIAMVLNMFVGMWLMG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKGSVERAREALAHMSSVFSGITLTK 1206
QY 1221 LPEGLVGLAKAQLOIIFPFRMLNLLTLGLLHGLVFLPVILSYVGPDPNPA 1272
Db 1207 FGGIVVLAFAKSIQIFFYFRMYLAMVLLGATHGLIFLPVLLSYIGPSVKA 1258

RESULT 5
US-11-191-274A-129
; Sequence 129, Application US/11191274A
; GENERAL INFORMATION:
; APPLICANT: Applera Corporation
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001536PROV
; CURRENT APPLICATION NUMBER: US/11/191,274A
; CURRENT FILING DATE: 2005-07-28
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 1278
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-191-274A-129

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLLWALLRLAQSPYTHIQPGYCAFDVDEG-----KNPELSGSLMTLSNVSCLS 59
Db 4 RGLAUGLLLLCPAQVFSQ-----SCVMGEGGIAGDKRYNCEYSG-----46
QY 60 NTPARKITGDHILILQKICPRLYTGTENTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKDLNLQPLQFLSRCPSCFYN 103
QY 120 FVNLHCHNTCSNQSLFINVTR-----VAOLGAGQLPAVVAYRAFYQHSFABOSYDSCSRV 175
Db 104 LNLFLCELACSPRQSQFLNVTATEDYDVPVNTQTKNKELOVYVQSPANAMYNACRDV 163
QY 176 RVPAAATLAVGTWCGVYGSALNQAQRWLNFGQDTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLGGKQADA-CNATNWIEYMFNKONGQAPFTTTPVFSDFPVH-----215
```

```
QY 229 VSGSQPLNEGVARCNESQGDVATCSCDCASCPAIARPO-----ALDSTFYLG 279
Db 216 ---GKEPMNATKGCDESDEVTAPCSCDCSVCGPKQPQPPPPATWITLGLDMVYIM 272
QY 280 QWPGSLVLIIILCSVPAVVTILL-----VGFVAPARDKSNVDPKGTSL 326
Db 273 WITTYMAFLVFPFAGFAVWCYKRYFVSYTPIDSDIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLQOFFCGMGTTWASWPLTILVLSIPVVALAAGLVTELTDTDPVLSWAP 386
Db 325 DPVSAAFEGCLARLFTRWGSPCVRNPGCVIPFSLVFITACSSGLVFRVTNPFVLSWAP 384
QY 387 NSQARSEKAFHQHGFPPFRTRNQVILTAPNRSYRVSLLLGPK-NFSGILDLLELL 445
Db 385 SSOARLEKEYFQHFQFPFRTEQLIIRAPLTKHIYQYPSPGADVFPFGPDLQILHQL 444
QY 446 ELQERLRHLQVMSPEAQRNLSQDICYAPLNPDNTSLYDCCINSLLQYQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVTILODCLAPLSPYNT---NCTILSVLNYFQNSHVLHKK 499
QY 506 NOTLMGQTSQVDWKDHFLYCANAPLTFKDGTAALALSCMADYGAPEVFPFLAIGYKGDYS 565
Db 500 GDDFP---VVADYHHTFLYCVRAPASLNDTSLHDECLGTFGGPFVFWLVLGGYDQNYN 556
QY 566 EBAALIMTSLNYPAGDPRLOAKLWEBAFLEEMRAFORFMAGMFMQVFTTFAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKNYKN---PNLTISFTAERSIEDEL 613
QY 626 NETTAEDLPATSYIVIVELYISLALGYSYSSWSRVWDSKATILGCGVAVLVGAVMAAMG 685
Db 614 NRESDSDFVTVVISVAIMFLYISLALGHYKSCRLLDVDSKVSGLGIAGILVILSVSACSLG 673
QY 686 FFSYLGIRSSVILQVPPVPLVSLVGADNIFIVLEYQRLPRRPGEBREVHIGRALCRVAP 745
Db 674 VFSYLGRLPLTVIEVIFPLVLAGVDNIFILVQAYQORDERLQGETLDQOGLRGVLEAP 733
QY 746 SMLLCSLSBAICFFLGAITPMPAVRTFALTGLAVILDFLLQMSAFVALLSDSKRQEAS 805
Db 734 SMFLSFSFSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLLOITCFVSLLLGLDIIKQKRN 793
QY 806 RLDVCCVKPQELPPGQ-GEGLLLGFFOKAYAPFLLHWITRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESLCFRFFKNSYSPLLKKDWMRPVIAIFVGLSPSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLPLNRYFEVGAUVYFVTLGYNFSSRAGMAICSSAG 924
Db 854 NKVDIGLDQSLSPDDSYMDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVGCGMG 913
QY 925 CNMFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDWLTP--SSCCRLYISGPNKDFPCPST 983
Db 914 CNMDSLVQOIFNAAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNMCSTT-MGSVRPSVEQFHKYLPWFILNDRPNIKCPKGGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR-CRPLTPBEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSSAVNLLGH 1029
QY 1041 DGOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMTVHTVLTQTSADFIDALKARLIASNTV-ETMGINGS--AIRVFPYSVFTV 1086
QY 1101 FYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLSRLNLLSIVMLVDTVGFMAWD 1160
Db 1087 FYEQYLTIIDDTIFNLGVSGLAIFLVTWLLGCELWSAVIMCATIAMVLNMFVGMWLMG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKGSVERAREALAHMSSVFSGITLTK 1206
QY 1221 LPEGLVGLAKAQLOIIFPFRMLNLLTLGLLHGLVFLPVILSYVGPDPNPA 1272
Db 1207 FGGIVVLAFAKSIQIFFYFRMYLAMVLLGATHGLIFLPVLLSYIGPSVKA 1258
```



```
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTTTPVSPDPVH-----215
QY 229 VSGGIQPLNEGVARNESQGDVATCQDCQCAASCPAIAAPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVATPCQDCSVCQPKPQPPPPAPWPTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSRWDPKGTSL 326
Db 273 WITYMAFLVFPGAFAVWCYKRYFVSYTIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGQGTWASWPLTILVLSVPPVVALAAGLVPTLTTPDVELWSAP 386
Db 325 DPVSAAFEGCLRLRFRWGSFCVRNPGCVIFPSLVFITACSSGLVFRVTTNPVDLWSAP 384
QY 387 NSQARSEKAFHDOHGFPPRTNQVILTAPNRSRYDLSLLGPK-NFSGILDLLLELL 445
Db 385 SSOARLEKEYFDQHFQGFPPRTTEQLIIRAPLTDKHIYQYPVPGADVPFGPPLDIQILHVL 444
QY 446 ELQERLRHLQVMSPEAQRNISLQDICYAPLNPDTNLSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVLNYFQNSHVLHDKK 499
QY 506 NOTLMGOTSQVDKDHFLYCANAPLTFKDGTLALSCMADYCAPVPPFLAIGYKGYKYS 565
Db 500 GDDFF--VYADYHTHFLYCVRAPASLNDTSLHDPCLGTFCGPVFPMLVGLDYDDQYN 556
QY 566 EBAELIMTFSNLNYPAGDPRLAQAKLWEAFLEEMRAFQRMAGMFQVTTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEKFTFNKYNKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEADLPATSYIVIFLYTSLALGSYSWSRVMSVDSKATLGLGGVAVLVGAVMAAG 685
Db 614 NRESDSVFTVVISAIMFLYISLALGHKSCRRLLVDSKSLGIAGILVLSSVACSIG 673
QY 686 FFSYLGISSVLQVPPFLVLSGADNIFIPVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTIVIEVPEFLAVGVDNIFILVQAYQDERLQGETDQOLGRVLGEVAP 733
QY 746 SMLLCSLSBAICFFGALTFMFAVTRFALTGLAVILDLLQMSAFVALLSDSKRQEAS 805
Db 734 SNFLSSFSEVTAFFGALSVMAVHTFSLFAGLAVFIDFLQITCFVSLLLGLDIKRQKN 793
QY 806 RLDDVCCVKBQLPPPGQ-GEGLLGGFQKAYAPFLHWTIRGVVLLFLFALFGVSLYSM 864
Db 794 RLDDIFCCVRGABDGTSSQASESCLFRFPKNSYSPILLKDWMPVIAIFVGVLSFSIAVL 853
QY 865 CHISVGLDELALPKDSYLLDFELFLNRYFEGVAPVYFTTILGYNFPSSAGHNAICSSAG 924
Db 854 NKVDIGLQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEBHDYTSKQGNMVGCGMG 913
QY 925 CNNEFSTOKIQVATEPPEQSXYLAIPASSWVDPDFDMLTP-SSCRLYISGPNKDFCPS 983
Db 914 CNNDLSUVQILFNAQLDNTVTRIGFAPSSWIDDYFQWVKPQSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCLNCSGIT-MGSVRPSVEQPHKLPFLNDPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPGKQRPQGGPMFLPFLNSDNPNPKCGKGAHAYSSAVNILLGH 1029
QY 1041 DQCVLASRFMAHYKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEPVFTYTNV 1100
Db 1030 GTRVGATYFTWYHTVTLQTSADFDLAKARLIASNTV-ETMGINGS--AYRVFPYSVYV 1086
QY 1101 FYEOYLITLPEGLPMLSLCLVTFPAVSCILGLDLBSGLNLLSIVMLIADTVGPMALWD 1160
Db 1087 FYEOYLITLDDTIFNLGVSGLGAIFLVTVMVLLGCELWSAIVMCATIAMVLNFMVGMWLMG 1146
QY 1161 ISYNAVSLNLNLSAVGMSVEFVSHITRSPASTKPTWLRARAKENATSMGSAVFAVAMTN 1220
Db 1147 ISLNAVSLNVLNWSGCISEVFCSHITRAFTVMKSGRVERAEALAHMGSSVFGITLTK 1206
QY 1221 LFGILVLGAQAQLIQIFFRNLNLITLGLLHGLVFLPVLISYVGPVNPA 1272
Db 1207 FGGIVVLAFAKSQIQIFVFRMYLAMVLLGATHGLIFLPLVLLSYIGSPSNKA 1258
```

RESULT 8

```
US-11-385-692-2401
; Sequence 2401, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385,692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2401
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2401
```

```
Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGMLLWALLRLAQSEPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
Db 4 RGLALGLLGLLLCPAQVFSQ-----SCVMYGECAIYGDKNYCNCEYSG-----46
QY 60 NTPARKITGDHLLILQKICPRLYTGPNTQACCSAKQVLSLEASLSITKALLTRCPACSDN 119
Db 47 ---PPRLPKDGVDDLQVELCPGFFFG-NVSLCCDVRQLTKDNLQLPLQFLRCRPSCFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFASQSYSCSRV 175
Db 104 LLNLFCELTCSPRQSOFLNVTATEDYVDPVTNQTNTNKLQYVYVQSQFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCGYVGSALCNAORWLNFGQDGTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTTTPVSPDPVH-----215
QY 229 VSGGIQPLNEGVARNESQGDVATCQDCQCAASCPAIAAPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVATPCQDCSVCQPKPQPPPPAPWPTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSRWDPKGTSL 326
Db 273 WITYMAFLVFPGAFAVWCYKRYFVSYTIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGQGTWASWPLTILVLSVPPVVALAAGLVPTLTTPDVELWSAP 386
Db 325 DPVSAAFEGCLRLRFRWGSFCVRNPGCVIFPSLVFITACSSGLVFRVTTNPVDLWSAP 384
QY 387 NSQARSEKAFHDOHGFPPRTNQVILTAPNRSRYDLSLLGPK-NFSGILDLLLELL 445
Db 385 SSOARLEKEYFDQHFQGFPPRTTEQLIIRAPLTDKHIYQYPVPGADVPFGPPLDIQILHVL 444
QY 446 ELQERLRHLQVMSPEAQRNISLQDICYAPLNPDTNLSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVLNYFQNSHVLHDKK 499
QY 506 NOTLMGOTSQVDKDHFLYCANAPLTFKDGTLALSCMADYCAPVPPFLAIGYKGYKYS 565
Db 500 GDDFF--VYADYHTHFLYCVRAPASLNDTSLHDPCLGTFCGPVFPMLVGLDYDDQYN 556
QY 566 EBAELIMTFSNLNYPAGDPRLAQAKLWEAFLEEMRAFQRMAGMFQVTTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEKFTFNKYNKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEADLPATSYIVIFLYTSLALGSYSWSRVMSVDSKATLGLGGVAVLVGAVMAAG 685
Db 614 NRESDSVFTVVISAIMFLYISLALGHKSCRRLLVDSKSLGIAGILVLSSVACSIG 673
QY 686 FFSYLGISSVLQVPPFLVLSGADNIFIPVLEYQRLPRRPGPREVHIGRALGRVAP 745
```

Db 674 VPSYGLPLTLIVIEVPLVLAVGVNDIFILVQAVQORDERLQGETLQQLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICFFLIGALTPMPAVTFTALTSGVLAVILDFLOMSAFVALLSDSRQAS 805
Db 734 SMFLSSFBETVAFPLGALSVMVHTFSLFAGLAVFIDFLQLQITCFVSLGLGDIKROEKN 793
Qy 806 RLDDVCCVKPQLPPPGQ--GEGLLGFFOKAYAPFLHWTIRGVLLFLFALFGVLSYM 864
Db 794 RLDFICCVGAEDGTSVQASECLFRFFKNSVSPILLKDWMPVIAIFVGVLSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDFLNLRYFEVAGPVYFTTLGYNFSSEAGMNAICSSAG 924
Db 854 NKVDIGLQSLMPDSDYMDYFKISQYLHAGPPVYFLERGHDTYSSKGQNMVCGMG 913
Qy 925 CNPFSTQKIQVATEFPEQSYLAIAPASSWDDFIDWLP--SSCCRLYISGPNKDFCPST 983
Db 914 CNDSLVQOIFNAALDNYTRIGFAPSSWIDDYFOWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLCKNCSMIT--MGSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDPACVR--CRPLTEGKQRPQGGDMRFLPNFLSDNPNKCKGGAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRPMAHYKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAPFVPTITNV 1100
Db 1030 GTRVGATYFTVHTVLTQTSADFDALKKARLIASNV--ETWINGS--AYRVFPYSVFV 1086
Qy 1101 FFEQVLTILPEGLFMLSCLVPTFAVSCLLIGLDRSLGILLNLSIWMILVDTVGFMAWD 1160
Db 1087 FFEQVLTITIDDTIFNLGVSGLAIFLVTWVLLGCELWSAVIMCATTIAMLVNMFVWLMG 1146
Qy 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFAITKPTWLERAKEATISMGSAVFAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGSISEFVCSHTITRAFTVMSKSRVERAEALAHMGSSVFSGITIK 1206
Qy 1221 LFGILVLGLAKAQLIQIFPRNLNLTITLGLHGLVFLPVILSYVGPDPNPA 1272
Db 1207 FGGIVVLAPAKSQIQIFVFRMYLAMVLIGATHGLIFLPLVLSYIGPSVUKA 1258

RESULT 9
US-11-301-094-2
; Sequence 2, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 1383
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-11-301-094-2

Query Match 19.4%; Score 1341.5; DB 6; Length 1383;
Best Local Similarity 27.5%; Pred. No. 2e-99;
Matches 369; Conservative 257; Mismatches 561; Indels 155; Gaps 40;

Qy 76 KICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSPNQSL 135
Db 63 EFCPHLLTGDN--KLCCPTSPAGBLTKQIAQARHILGRCPSCFDNFAKLMCEFTCSPNQD 121
Qy 136 FTVNTRVAOL--GAGOLPAVVAEAF----YQHS--FAEQSYDSCSRVRVPAATLAVG 186
Db 122 FVISEMKFIEKEGFTPEYQAEAVVTVFRLSTDPAGMPSCKDVTFGQPALRV- 180
Qy 187 TWCGVVGSAALCAQNRWLNFGQDTGNGL--APLDITFHLLEPGQAVSGIOP--LNEGVARCN 244

Db 181 -WC---TSTPCTLTNWLEFIGNLQNLNLNPIHTIKFLLYDPIKTPPSDRSTYMNVTGCD 236
Qy 245 ESQDDVATCSQDCQCAASCPAIARPOALDSTFYLGOM-----PGSLVLIILICSPAVVT 299
Db 237 KSARVGMVACSTSEC--NKEBYANLIDLDGKTSQTCNVHGIACINIFVMLAFIGSLAV 294
Qy 300 ILLVGF-----RVAPARKSKMVDPKKGTSLSDKLSFSFTHLLGQFF 341
Db 295 LLCVGFVTSYDEDTNLRQTSGEESPRNRK-----RTGANI-----HNFMENNA 342
Qy 342 QGWGTWASWFLTILVLSVIPVVALAAGLVFTELTDPELWSPASQARSKAFHQHF 401
Db 343 RDIGMAGRNPKSHFFIGCAVLIFCLPGMIYHKESTNVDMWSSPRSRARQSEMVFNANF 402
Qy 402 GPPFTNQVILTAPNRRSSRYRDSLLLGPNKPSGILDLLELLELLELQELRLHLQWSPRA 461
Db 403 GRPORYQOIMLL--SHRDFQSSGKLYG-----VFHKDIFEEFLDILNAIKNISTQDSG 455
Qy 462 ORNLSLODICYAPLNPDNTSLYDCCINSLLOVFQNNRTLLLLLTANQT-----L 509
Db 456 -RTITLDDVCYRPMGPG---YDCLIMSPYFQGNKEHLDKMSKEETSEDDDAFYF 510
Qy 510 NGQTSQVDWKDHFYLCANAPLTFKDGATALASCMDAYGAPVPPFLAIGYKGYSEABA 569
Db 511 SSEATTDSEMNHMAACIDQPMQK--TKSGLSCMGTYGPGSAPNM--VFGKSTNHOAANS 567
Qy 570 LIMFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTFTAERSLEDEINRT 629
Db 568 IMMTILVTQ--RTEPEIQABELWEKEFLKCKEYREKSPKVI--FSFMAERSTIDBIENDA 624
Qy 630 AEDLPPIFATSYIVIFLYISLALGSY-----SSWSRVMDVSKATLGLGVAVVVLGAVMAAM 684
Db 625 KDEIVTVVIALAFIYGYTFSYGRVFCENQLMS--ILVHSRCLGLMSVILNLLSFCWS 683
Qy 685 GFFSYLGRSSILVQLVQVVPFLVLSVGADNIFIFVLEY--ORLP-----RRGPBRVHIGRA 739
Db 684 GIFSMFGLHPVKNALVQGFVVTLIGVCRTEWVKYQAQRVSMFYMSPDQCPEI--VGMV 742
Qy 740 LGRVAPSMLLCSLSEACIFPLGALTPMPAVTFTALTSGVLAVILDFLOMSAFVALLSDS 799
Db 743 MAGTMPAMFSSSLGCASFFFIGGFTDLPAIRTFCLYAGLAVILDVVLHCTITFLALFVWD 802
Qy 800 KRQASRLDVCCVKPQLPPPGQEGEGL-----LLG-----FFQKAYAPFLHMI 844
Db 803 QRELNG-----KP-EFFPYQIKLLGAYLIGQRATDFTMTQFFHFQVAPFLMHRM 853
Qy 845 TRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDFLNLRYFEVAGPVYFT 904
Db 854 TRIITGIIFIASFITTVILSSKISVGFQDSMAFTEKSYISTHFRYLDKFPDVGPPVFTV 913
Qy 905 TLGYNFSSEAGMNAICSSAGCNPFSTQKIQVATEFPEQSYLAIAPASSWDDFIDWLT-P 963
Db 914 DGELDWHRPDVQNKPFCTPFGCSDTSFGNIMYAVGHTQTLYLSGEMYNIDWILEMISK 973
Qy 964 SSCRLYISGPNKDFCPSTVNSL-----NCLKNCH-----SITMGSV---RPSVEQPH 1009
Db 974 SPCCVKYVHDN--TFCTNRNKSALDDKACRTCMDFDYVANSYKSSIMYHRPSLEVY 1031
Qy 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQDQVLASRFMAYHKLPL--KNQDYTEAR 1067
Db 1032 RHLRHFLEDTNSECFCVGRASFDAISFTSRGRIQASQFMTHFKKLSISNSDFTKAMD 1091
Qy 1068 ABELAANITADLRKVPCTDPAPFVPTITNVFVEQVLTILPEGLFMLSCLVPTFAVS 1127
Db 1092 TARMVSRRLERSI-----DQTAHVFAYSKIPFPFPEQYSTIMPILTQTFLITVGVFGII 1145
Qy 1128 CILLGLDRSLGILLNLSIWMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITR 1187
Db 1146 CVTLGIDYKGAACAVICQVSNFYHIVAFWYIFNIPVNALSATNLVMSGILLIEFSVNLK 1205
Qy 1188 SPAISTKTWLERAKEATISMGSAVFAGVAMTNLFGILVLGLAKAQLIQIFPRNLNLT 1247
Db 1206 GYACSLRQAKDRABSTVGSIGPIILSGFVVTMAGSTMFLSGAHLQIITVYFKFLIT 1265


```
Best Local Similarity 38.4%; Pred. No. 7.2e-55;
Matches 164; Conservative 88; Mismatches 138; Indels 37; Gaps 10;

QY 863 SMC-HISVGLDELALPKDSYLLDYFLNRYFVGVAPYVFTTLGYNPSSA-GMNAIC 920
Db 2 ALCTRIEAGLEOQIALPRDSYLLQGYFNSIYSEYLRVGPPLYFVVK-DYNYLSKHTNQLC 60
QY 921 SAGCNPFSTKIOVATEFQSQSYLAIPASSWVDDFIDWLTPS--SCRLYISGNKDK 978
Db 61 SISHCDNSLLNEISRASLVPTSSYIAKPAASWLDLFLWISPEAFSCCKP-----TND 116
QY 979 FCPSTVNSLNCNKVMSITGVS-----RPSVEQFHKPLWFLNDRPMI 1022
Db 117 YCPDDPPCCLPDEGPGCGKCKDCTTCFRHSDLVNDRPSTAQREKLPWFLDALPSA 176
QY 1023 KCPKGLAAYSTSVNLT--SDQVLAIRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 177 DCAKGHGAYTNSVDLNGYEGGVIQASEPRTYHTPLNRQGDYVNAIRARDPSAIISSSL 236
QY 1081 RKVPGTDPAFEPYPTITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 237 K-----MDIPFSVVFIFFEQYLDIWKLLINITVALGAIFV-CUIITSWSSAI 287
QY 1141 NLLSIWMLVDVGVFMAWDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
Db 288 LLLVLMIMILDMGVMAILGIQLNAVSVVNLMSIGIAVEFCVHIVAFMVSIGDR-SQR 346
QY 1201 AKERTSMGSAVPAGVAMTNLPGILVLGLAKAQLQIIPFRNLNLIITLGLHLGVFLPV 1260
Db 347 AKTALCTMGASVFSGITLTKLVGLVLCFSTSEIFVYVYFOMYLAIVITGLHLGVFLPV 406
QY 1261 ILSYVGP 1267
Db 407 VLSLFGP 413

RESULT 12
US-10-461-673-16739
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Wang, Jian-Rui
APPLICANT: Ren, Felyan
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Xue, Aidong
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Zhi Wei
APPLICANT: Yang, Yonghong
APPLICANT: Goodrich, Kyle W.
APPLICANT: Wang, Dunrui
APPLICANT: Yamazaki, Victoria
APPLICANT: Ujwal, Manusha L.
APPLICANT: Ma, Yunging
APPLICANT: Chen, Rui-Hong
APPLICANT: Ghosh, Malabika
APPLICANT: Weng, Gezhi
APPLICANT: Haley-Vicente, Dana
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 823
CURRENT APPLICATION NUMBER: US/10/461,673
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: PCT/US02/29964
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 10/245,014
PRIOR FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/323,739
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: PCT/US02/29636
```

```
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt FL_genes Version 6.0
; SEQ ID NO 16739
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-673-16739

Query Match 8.5%; Score 588.5; DB 6; Length 967;
Best Local Similarity 21.3%; Pred. No. 1.4e-38;
Matches 216; Conservative 192; Mismatches 359; Indels 249; Gaps 37;

QY 334 HT-----LLGQFFQGWGTWASWPLTILVSVIPVVALAAGLVF--TELTDPVLSAP 386
Db 115 HTDCEGLSLRTFOWLQWGAHPWIFLLAPLMLTAALGTGLYLPKDBEEDLEBHYTPV 174
QY 387 NSQARSEKAFDQHFGPPFTNOVILTAPNRS-----YRYSLLGPKMPSGIL 436
Db 175 GSPAKAERRFQGH-----FTTNDSYRFSASRRSTANFVSLVSVSLSLDPATFAEVS 230
QY 437 DLDLLELLELQERLRLHLQVMSPEAQRNISLQDIC--YAPLNPDNTSLYDCI--NSLLQ 492
Db 231 KLDGAVQDLRV-----AREKGSQIQVQVCARYAL-----CVPNPILY 270
QY 493 YFQNRITLLLTANQTLMGQTQVQDWKDHFLYCANAPITFKDGTALALSCWADYGA-PVF 551
Db 271 AWQVKNKTLNL-----SSISFPAY-----NHRGHPLY 296
QY 552 PFLAIGGY-----KGKDYSEAEALIMTFSLNNYPAGDPRL-AQAKLWEAEFLSEMRAP 603
Db 297 LTGPFGGYILGSLGMLGQALLRAKAWRLLYLK--TEDPEYDVQSKQWLHLLDQFTNI 353
QY 604 QRRMA-----GMFQVTFTAERLSLEDEINRTAEDLPFATSYIVIFYISL 649
Db 354 KNILALKKIEVPGVGLQGQEKVHPFTSLRQLEFEATSVTVIPVFLAYLILFAVT 413
QY 650 ALGYSYSSWSRVVDSKATLGLGGVAVLGVAMAMGPFYSYLGIRSLVILQVVPFLVSV 709
Db 414 SCFRPD-----CIRNKMCAAFGVISAFLVAVSGFLLHLHIGV-PFVIIVANSPLILGV 467
QY 710 GADNIFIFVLEYQRLPRPGEPRVHIGRALGRVAPSMLLCSLSAEICFFLGALTPMPAP 769
Db 468 GVDDMFIMISAWHK-TNLAGDIR-RMSNVYSKAAVSITITINILALYTGIMSSFRSV 525
QY 770 RTFALTSLGAVILDFLLQMSAPVALLSDSKRQEAASRLDVCCKVPQELP-----819
Db 526 QCFCIYTGTTLLFCYFNITCFGAPMALDGREV-----VCLCWLKKAADPKWPSKKFC 580
QY 820 -----PFGQEGEL--LLGFFQKAVAPFLHLHWITRGVLLFLALFGVSLVSMCHISVGL 871
Db 581 PFGSVDPDEHGTDHPMSLFFRDYFGPFLTSESKYFVFIYLVLYIIISYGCFFHQEGL 640
QY 872 DQELALPKDSYLLDYFLFLNRYF-EVGAIVYVFTTLGYNFSSSEAGMNAICSSAGCNPF 930
Db 641 DLRLNASDSDSYITPYFNVEENVFSDYGPVMVITVKVDYWDK-----DV 685
QY 931 TQKIQYATEFPEQSYLAIP--ASSWVDDFDWLTPSSCCRLYISGNKDKKPCPSTVNSLN 988
Db 686 RQKLENCIKIFKNVYVDKNLTFEFLDAYVQ-----YLGKNSQD---PNEKNT--730
```



```
QY 989 CLKNCMSITWGSVRPSVEQPHKYLPMFLNDRNENIKCPKGLAAYSTSVNLTSDGQVLASR 1048
Db 731 -----FMNIPDPLGNFN-----FQHDINISSNEIISR 761
QY 1049 FMAYHKPLANSQDYTEALRAARELANITADLRKVPDTPAEPVPPYTTITNVPYQOYLT 1108
Db 762 GFIOITDVS-----AKKI-----LLFQURRI--AEDCQIPLMVYNQAFYFQYAAI 809
QY 1109 LPE-----GLFMLSICLVP-----TFAVSCILLGLDLRLSLNLLSIVMILV 1150
Db 810 LEDTVRNVLVASAMFIVSLLIPIYPLCSLWTFAGISVIVG-----852
QY 1151 DTGFMALWDISYNVSLNLSAVGMSVEFVSHITRSPAISTKPTWMLERAKEATISMG 1210
Db 853 ---TGFMAPKVNLDISMINLVICIGFDFSAHISYAFVSSQPSVQNSQKSEALYLLGY 910
QY 1211 AVPAGVAMTNLCILVLGLAKQOLIQIFPERLNLITLGLLHGLVPLVILSYVG 1266
Db 911 PVLQS--AISTIIIGVCVLAARAKAYIFRT--PFKIMFLVMIFGAHGLIPIPVLTFFG 964

RESULT 13
US-11-332-764-2
; Sequence 2, Application US/11332764
; GENERAL INFORMATION:
; APPLICANT: Wisotzkey, Robert G.
; TITLE OF INVENTION: PATCHED HOMOLOG 2 (PTCH-2) DISRUPTIONS, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: RELATING THEREO
; FILE REFERENCE: R1673 CIP/75658.064000
; CURRENT APPLICATION NUMBER: US/11/332,764
; CURRENT FILING DATE: 2006-01-13
; PRIOR APPLICATION NUMBER: US 60/413,543
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 10/669,143
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-332-764-2

Query Match 8.4%; Score 583.5; DB 6; Length 1182;
Best Local Similarity 24.2%; Pred. No. 4.8e-38;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SPSHTTLLGQ-----FPQG-----WGTWVASWPLTILVLSVIPVVALAAGLVTELT 376
Db 20 SSAPHILAGSLQAPLWLRAYFQQLPSLCRIQKHCQKVLFLGLVAFGALALGLRVAVIE 79
QY 377 TDPVELWASPNQARSEKAFHQHFG--PFRITNOVILTAPNRSYRSDSLILGPNKFSGI 435
Db 80 TDLEQLWVEVGRSVQELHYTKEKLEBAAYTSQMLI-----QTAHQEGGNVLTPE---A 131
QY 436 LDLDLLELLELLEQLERLHQVSPRAQNISLDICV---APLNDN-----TSLYDCC 486
Db 132 LD-----LHQALATASKVQSVLYGKSWDLNLCIKYSGVPLIENGMIERIEKLFPCV 184
QY 487 INSLLOYFQNNRTLLLTANQTLMGQTSQVMDK---HFLYCANAPLT-----FKDGTALA 539
Db 185 ILTPLDCFWGAK---LQGSAYLPGRPDIQWTNLDPPQLLELGPFPASLEGFRELDKA 241
QY 540 LSMADYGAAPV-----PFLA---IGYKGDYS-----565
Db 242 QVGQAYVGRPCLDPPDPHCPSPAPNHRSPAQENVAQELSGGCGFSGHKFMHQBELLGG 301
QY 566 -----EAEALMTF-----SLNNYPAGDPRLAQAKLWEARFLEMAFORRMAG 609
Db 302 TARDLOGQLLRALQSTFLMLSPROLYEHFRGDYQTHDIGWSEQASWVQAWQRFRVQ 361
QY 610 MFQVTFABRSLE-----DEINRTTAEDLPIPATSYIV--IPLYISLAGSYSSWS 658
```

```
Db 362 LAQEAALPANASQOIHAFPSSTTLLDDILRAFSE-----VSTRVVGGYLLMLAYACVTMLRWD 417
QY 659 RVWVDSKATLGGVAVVILGAVMAAGFFSYIGIRSSILVILQVVPFLVLSVGADNIFIFV 718
Db 418 ---CAQSOAGVLAGVLLVALAVASGLCALGIGTFNAATTQVLPFLALGIGVDDIFLLA 475
QY 719 LEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTMPMPAVRTALTSL 778
Db 476 HAFTRAP--PDTPLPERMGECLRTGTSTVALTSVNNMVAFMAALVPIPALRAFSLOAAI 533
QY 779 AVILDFLLQMSAFVALLSLDSKQASRLDVCCC-----VKPQE-----817
Db 534 VVGCNFAAVMLVFPAILSLDLRRHRQRDLVLCFSPSCSAQVIOQLPOELGDRVAVGI 593
QY 818 -----LPP-----PG-----QEGE--826
Db 594 AHLTATVQAFTHCEASSQHVVITLPPQAHLLSPASDPLGSELYSPGSTRDLSQEGTG 653
QY 827 -----LLLGFFQKAYAPFLHWHITRGVLLFLALFGLVSLYSMCHISVGLDQ 873
Db 654 POACRPELLCAHWTLAHFARYQFAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713
QY 874 ELALPKDSVILDDYFLFLNRYPEVGAPVYFVTTLGYNFS--SEAGMNAICSSAGCNPFSTQ 932
Db 714 TDVVPRTGKEHAPLSAQLRYFSL--YEVVALVTQGGFDYAHSORAL-----FDLHQ 761
QY 933 KIOYATEPPEQSYLAIPAS-----SWVDDF-----IDMLTPSSCCRLYISGPNK 976
Db 762 RFSSIL-----KAVLPPTATQARTWLHYRSLWLOGIQAAFDQDQWASGRITCHSYNGSED 816
QY 977 D-----KFCPSTVNSLCLNKMCSITM-----GSVRPSVEQPHKYLPMFLNDRNENIKCPK 1026
Db 817 GALAVKLLIQTCNAQEPDLDFSLTTRKLVDKBGLIPP--ELFVMTGLTVWVSSDPL-----869
QY 1027 GGLAR-----YST---SVNLTSQOVLASRPMAYHKPLKNSQDYTEALR 1067
Db 870 -GLAASQANFYPPPELWLDKDYDTTGENLRIPAAQPLEPAQFPFLHLGLQKTADFVEAIE 928
QY 1068 AARELAANI--TADLRKVPDTPAPEVPVYITNVFVQYVLTILPEGLFMLSIC--LVPTF 1124
Db 929 GARAACTEAGQAGVHAYPSGSPF-----LFWQYVGLI--RRCFLAVCILLVCTF 976
QY 1125 AVSCILLGLDLRSGLLNLISIVMILVDTVGFMALWDISYNVSLNLSAVGMSVEFVSH 1184
Db 977 LVCALLLSPWTAGLI--VLVLAMTVELFGINGFLGKLSAIPVWLIVASIGIGVEFTVH 1035
QY 1185 ITRSPAISTKPTWMLERAKEATISMGSAVPAGV---AMTWLPGILVLGLAKAQLIQIFPFR 1241
Db 1036 VALGFELTSHGSRNLRAA-----SALEQTAPVPTDGAUSTLLGLLMLAGSNPDFIIRYFFV 1090
QY 1242 LNLITLLGLLHGLVPLVILSYVG 1267
Db 1091 VUTVITLLGLLHGLLPLVILSLGP 1116
```

```
RESULT 14
US-11-337-244-149
; Sequence 149, Application US/11337244
; GENERAL INFORMATION:
; APPLICANT: William Matthews
; APPLICANT: Mark Moore
; APPLICANT: Russell Phillips
; APPLICANT: Michael V. Wiles
; APPLICANT: Thadd C. Reeder
; APPLICANT: Robert G. Wisotzkey
; APPLICANT: Keith D. Allen
; APPLICANT: Helen Baribault
; APPLICANT: Thomas J. Brennan
; APPLICANT: Catherine Guenther
; APPLICANT: Robert Klein
; APPLICANT: Christopher J. Kirk
; APPLICANT: Simon X. Xie
```


Db 997 IAAVLCSGPTNYN--CDGRLCKMKWNEV-----INP--EGFVNYLTGWFVD--NM 1044
QY 1023 KCPKGLAAYSTSVNLTSDQVLASRFMAYHKPLKNSQ-----DYTEALRAARELA 1073
Db 1045 MYVQSASFPTPPGMEYNEKL--AKVWPAEPLLYSQMPFYQNDLIDTPAIVKMIER 1102
QY 1074 ANITADLRKVPCTDPAFEVFPYTTITNVFVEQVLTILPEGLMLSLCLVPTPAVSCLLGL 1133
Db 1103 ATCE-----EYSEGLSNHPGIAFTFWEQVLT-LRNWLFQ-AICII-ALAVFCVISIL 1153
QY 1134 DLRSGLLNLSVIML--VDTVGFMALMDISYNAVSLINLSAVGMSVEFVSHITRSFAI 1191
Db 1154 MFNPAATLIMCIVITITIELGFGMLGMIKKNPISAVTLICAVGIGVEFTAHVELAF-L 1212
QY 1192 STKPTWLERAKGATISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIFPFRMLNLLTLGL 1251
Db 1213 TALGTIDQRLSCLOHMFVYVYHGAISTFL-GVNMVLVSEFDFVYVYFTYTWLTLLVALGV 1271
QY 1252 LHGLVFLPVILSVGPD-----VNPALALEQKAEBA 1283
Db 1272 FNGLCVLPVILTVGPKBELTTDGSVLPFPPLPRQQAES 1314

RESULT 17
US-11-360-355-133805
; Sequence 133805, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 133805
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (794)...(794)
; OTHER INFORMATION: xaa can be any naturally occurring amino acid
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=seqid 60228; Strand=-; Positions=1
; OTHER INFORMATION: -173,414-469,516-626,687-921,1040-1216,1282-1406,1470-1579,1627
; OTHER INFORMATION: -1785,1833-1933,1984-2145,2192-2241,2295-2457,2503-2695,2750
; OTHER INFORMATION: -2985,3164-3208
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_496761.1; Match level="QueryCoverage
; OTHER INFORMATION: =99%, HitCoverage=86%, E-value=0.0, Identity=48%"; Hit descriptio
; OTHER INFORMATION: =Patched Related (90.8 kD) (ptr-18) [Caenorhabditis elegans]
; OTHER INFORMATION: pir||T26683 hypothetical protein Y38FLA.3 - Cae
US-11-360-355-133805
Query Match 6.3%; Score 437; DB 7; Length 831;
Best Local Similarity 22.5%; Pred. No. 2.2e-26;
Matches 189; Conserved 139; Mismatches 303; Indels 210; Gaps 30;
QY 590 KLEWEAFLEEMAFQRMAGMQVTFATRSLEDEINRTAEDLPFTATSYIVIFLYISL 649
Db 21 KKWEHSVYN--FAMNSRGDPLKLVHTSEGLVSEVRRTGIEVLPLMPLISLVILFTVI 78
QY 650 A-----LGSY-----SSSRVSVYDVKATLGLGGVAVVLGAVNMAAGFFSY 689

Db 79 TSLKRDQIRSKPWEALVGFCTVTITSLKRDQIRSKPWEALVGFCTPILSVGASGFTLFW 138
QY 690 LGIR-----SSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGEPRVHIGRALGRV 743
Db 139 LGFEFLPGFBFLPLVVPFLIILAIGVDDVFIPLHCWAHTD--ESKOLRERVADLLGSA 196
QY 744 APSMLLCSLSBAICFLGALTTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLDSKROE 803
Db 197 GFSVTITSLTNWLSSTIGATFTPAIRFCFLFISVAVLYAYLYQLFFYTAVMVIGAQRBA 256
QY 804 ASRLDVCCCVK--POELPPPGQ-----GEGLLGFGFKAYAPFLPLHW 843
Db 257 DERNAYLFCIKRVPKRSKPVATLKETNWLRLGSTNLMLRGLSDLV-----DRYVDVFMVSW 311
QY 844 ITR-----GVVLLFLAL--FGVSLYSWMCHISVGLDQELALPKDSYLLDFLFLNR--YFEV 896
Db 312 TARILLALSLIYYAFSAYGVAQIKVAQIKVGLTSEKFLDSDSPLELVLRLQNNVIFKEG 371
QY 897 GAPVVFVTTLVGNFSSSEAGMNAICSSAG-CNNFSFTQKIYATEFPEQSYLAIPASS--- 952
Db 372 QOMAVFVNSPG---DLRKPMAVFNSPGDLRKPSPAIPIKMLRLEFEQSGNSGVGSSTQM 428
QY 953 WYDDFIDWLTSPSSCCRLYISGPNKDFCFTSVNSLNCMLKNCMSITMGSVR-----P 1003
Db 429 WLNTVLPFIG-----LQNRGSIINRGSIDPRYKLYDFF 461
QY 1004 SVEQHKYLPWF-----LNDRNLIKCPKGLAAYSTSVNLTSDQGVLASRFMAYHKP 1055
Db 462 SIPEYHRWSHFVSLGPKEDCLNERP-----SCINKPFDCLNERPSCINKFP-FSTG 511
QY 1056 LKNS--QDYTEALRAARELAA-----NITADLRKVPCTDPAFEVFPYTTITNVFVEQVLT 1107
Db 512 FQNAVMSDRVLVLRWQRLAADYSQLNLT-----VYEDFSMVADQMYADQLLS 560
QY 1108 I-----LPEGLFMLSCLVPTFAVSCLLGL-----DLRSGLLNLSIVMILVDTVTGFMAL 1158
Db 561 IPSVTIQTVAFALLCMTFVAFALLCMTFVLMFTPSISTILPGTACVLSINLGVFGLLFY 620
QY 1159 WDISYNAVSLINLSAVGMSVEFVSHI-----TRSFATSTKPTWL----- 1198
Db 621 WSIDLDPISTMTTLTMAIGLSVDFVAHISFHYHYKGMTTLTMAIGLSVDFVAHISFHYHYKGE 680
QY 1199 -----ERAKEATISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIF----- 1238
Db 681 FDSRERLRHALSSIANPMLQ-AALSTVLSMLVLIHRYMVQVFKVVLVVLVGLVHG 739
QY 1239 ---FPRLNLLITLLGLHLGFLPVLISYV-----GPDVNP-----ALALEQKRA 1280
Db 740 LVVFKVVLVVLVGLVHGLVVLVVYAAIPFQKSGATAKPKVPLNRRRTAVTXQKRRR 799
QY 1281 E 1281
Db 800 E 800

RESULT 18
US-10-461-673-16750
; Sequence 16750, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.

APPLICANT: Wang, Dunrui
APPLICANT: Yamazaki, Victoria
APPLICANT: Ujwal, Manuhea L.
APPLICANT: Ma, Yunging
APPLICANT: Chen, Rui-Hong
APPLICANT: Ghosh, Malabika
APPLICANT: Wang, Gezhi
APPLICANT: Haley-Vicente, Dana
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 823
CURRENT APPLICATION NUMBER: US/10/461,673
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: PCT/US02/29964
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 10/245,014
PRIOR FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/323,739
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: PCT/US02/29636
PRIOR FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: US 10/245,817
PRIOR FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/323,349
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: PCT/US02/29001
PRIOR FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: US 10/243,552
PRIOR FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: US 60/322,511
PRIOR FILING DATE: 2001-09-13
PRIOR APPLICATION NUMBER: PCT/US02/25485
PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 17116
SOFTWARE: pt_FL_genes Version 6.0
SEQ ID NO 16750
LENGTH: 783
TYPE: PRT
ORGANISM: Homo sapiens
US-10-461-673-16750

Query Match 51.1%; Score 349; DB 6; Length 783;
Best Local Similarity 21.5%; Pred. No. 2.7e-19;
Matches 193; Conservative 160; Mismatches 335; Indels 208; Gaps 42;

QY 444 LLEQLERHLQWSEAPQNSLQDICYAPLNPNTSLYDCINSLLQYFON----NRT 499
DB 10 ILKJHAAVTKIQV--PRPGFNFTFAHICI--LNNDKTCIADDIVH-VLEELKVARATNRT 64
QY 500 LLLLTANQTLMGQTSQVDWKDHFLYCANAPLT-FKDGTAALSCMADYGAPVFPFLAIGG 558
DB 65 NFAIT-----YPITHLKDGRA-----VYNGHQLGG 89
QY 559 --YKGD-YSEALIMTSLNNYPAGDPRLAQAKLWEAPLEEMRAFQRMAGMFQVTF 615
DB 90 VTVHSKDRVKSAEAIQLTYLQINSLNDMAER--WESSPCDITVRLFKSNSKVKMYPY 147
QY 616 TAERSLEDEINTTAEDLPIFATSIYIVPLYSLSALGSSYSSRWVWDSKATILGLGGAV 675
DB 148 TSS-SLRDFQKTSRVSERYLVTSLIV--VTMAILCCS--MQDCVRSKPLWGLGLVT 201
QY 676 VLGVMAAGMFTSYLGIRSLVILQVPPVLVSVGADNFIIFVLEYORLPRPGPREVH 735
DB 202 ISLATITAGIINLTGCKYNSTFLG-VPFVMLGHGLYGTFFMLSSW----RTREDQHVK 256
QY 736 IGRALGRVAPSMLLCSLSA---ICFFLGA--LTPMPAVRTFALTSGLAVIDFLQMSA 790
DB 257 -ERTRAVYADSMLSFLTAMYLVTFGIGASPPTNIEAARIFCCNSCIAIFENYLVLSF 315
QY 791 FVALLSLDSKREARSLDVCCVKP-----QELP-----PRQGE----- 825
DB 316 YGSSLVFTGYIENNYQHSIFCRKVPKPEALQKPAWYRFLLTARFSEDTAEGEEANTYBS 375

QY 826 GLLLGFFQKAYAPFLHLHWITRGVILLFLALFGVSLYSMCHISVGLDQBLALPKDSYLLD 885
DB 376 HLLVCLKRYCYCDWITNTVVKPVLFYLIYISFALMGVLYQVSGSDLSNIVATATQTTIE 435
QY 886 YPLFLNRYPEVGAPVYFVTTLGYNSSSEAGMAICSSACGNPFSFTQKIQYATEPEQSY 945
DB 436 YTTAQOQYFSNYSPIV-----GFYIYESIEYWNYSVQEDV 470
QY 946 LAIPAS-----SWDDPFDMLTPSSCCRLVIS--GPNKDKFCPSVNSLNCNKCMSITMGS 1000
DB 471 LEYTKGFVRISWFSYLVLR-----KLVNVTGLPKCNTDMLRNSP--LK----- 514
QY 1001 VRPSVEQPHKYLWFLNDRPNIKCPKGGIAAYSTSVNLTSQGVLASRFMAYHKPLK-NS 1059
DB 515 -----APQFSHQEDIIFSK-----KYNDEVD-----VVASRMFLVAKTMEINR 553
QY 1060 QDYTEALRAARELANITADLRKVGTDPAPEVPFYTITNVFEYQVLTILPEGLFMLSIC 1119
DB 554 EELYDLLETLRRL--SVTSKVFIV-FNPSP-----VYMDRYASSL--GAPLHNSC 599
QY 1120 LVPTFAV--SCULLGLDLRSGLNL--LSIVMILVDTVGFMALWDISYNAVSLINLVA 1174
DB 600 ISALFLLFSAFLVA-----DSLINWVITLVVSVEGVIGFTLWKVELDCISVLCIIVG 655
QY 1175 VGMSEVFSVSHITRSPAIS---TKPTWLERAKBATISMSGSAVFAGVAMTN-----LPGILV 1226
DB 656 INVTIDNCAPMLSTFVLGKDFTRTKWKNALV-----HGVAILOSYLCYIVGLIP 706
QY 1227 IGLAKAQLIQIFPFLNLLITLLGLLHGLVFLPVILSVYGVDPVNPALALEQKRAE 1282
DB 707 LAAVPSNL-TCTLFRCLFLIAFVTFPHCFAILFVILTFPPS-----KKRKEK 754

RESULT 19
US-11-360-355-141351
Sequence 141351, Application US/11360355
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey
APPLICANT: Du, Zijiang
APPLICANT: Guo, Liang
APPLICANT: Kovalic, David
APPLICANT: Lu, Maolong
APPLICANT: McCarter, James
APPLICANT: Miller, Nancy
APPLICANT: Williams, Deryck
APPLICANT: Vaudin, Mark
APPLICANT: Wu, Wei
TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
FILE REFERENCE: 38-21(53885)
CURRENT APPLICATION NUMBER: US/11/360,355
CURRENT FILING DATE: 2006-02-24
NUMBER OF SEQ ID NOS: 171306
SEQ ID NO 141351
LENGTH: 204
TYPE: PRT
ORGANISM: Heterodera glycines
FEATURE:
OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 67774; Strand=+; Position=1
OTHER INFORMATION: -32,241-399,575-746,940-1023,1087-1203,3133-3154
FEATURE:
OTHER INFORMATION: Homolog annotation: Hit ID=NP_999487.1; Match level="QueryCovera
OTHER INFORMATION: =97%, HitCoverage=14%, E-value=2e-44, Identity=468"; Hit descrip
OTHER INFORMATION: =Niemann-Pick C disease protein [Sus scrofa] gb|AAD47090.1| Niem
OTHER INFORMATION: -Pick C disease protein [Sus scrofa] sp
US-11-360-355-141351

Query Match 4.7%; Score 325; DB 7; Length 204;
Best Local Similarity 39.6%; Pred. No. 3.1e-18;
Matches 76; Conservative 35; Mismatches 65; Indels 16; Gaps 3;
QY 585 RLAAQAKLWEAPLEEMRAFQRMAGMFQVTFTAERSLEDEINTTAEDLPIFATSIYIVIF 644

RESULT 22
US-11-360-355-141367
; Sequence 141367, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing

RESULT 23
US-11-214-063A-1670
; Sequence 1670, Application US/11214063A
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: KAIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE

```

; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/11/214,063A
; CURRENT FILING DATE: 2005-08-30
; PRIOR APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1670
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-11-214-063A-1670

Query Match      3.3%; Score 225; DB 6; Length 542;
Best Local Similarity 20.4%; Pred. No. 1.8e-09;
Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;

QY 744 APSMLCSEA---ICFPGA--LTPMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSID 798
Db 23 ADSMLSFSLTAMYLVTFGIGASPTNIEARIFCCNSCIAIFFNVLVYVSGSLVPT 82
QY 799 SKRQEARSLDVCCVKP-----QELP-----PPQGE-----GLLGRFQ 833
Db 83 GVIENNYQHSIFCRKVPKPEALQEPAMYRFLLTARFSDTAEGBEANTYESHLLVCF 142
QY 834 KAYAPFLHWITRGVLLFLALFGVSLYSMSCHISVGLDOELALPKDSYLLDYFLNRY 893
Db 143 RYCDWITWYKPVVLYLYISFALMGYLOVSEGSLSNIVATATQITBTYTAQKY 202
QY 894 FEVGAPVYVFTLLGVNFSSEAGNACSSAGCNSFTQIOYATFEPPQSYLAIPAS-- 951
Db 203 FSNYSPI-----GFVYESIEYMTSVQEDVLEVTGKV 237
QY 952 --SWDDFTDMLTPSSCCBLYIS-GENKDKFCPESTVNSLNCNKMSITMGSVRSVQF 1008
Db 238 RISWFESYLYLR-----KLVSTGLPKQNTDMLRNSF--LK----- 273
QY 1009 HKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLK-NSQDYTEALR 1067
Db 274 ---APQSFHQEDIIFSK---KYNDEVD-----VVASRMFLVAKTMEINREELYDLLE 320
QY 1068 ABELAANTADLRKVPGTDPAFEPVPTITWVFEQYITLILPEGLFMLSCLCLVPTFAV- 1126
Db 321 TLRL--SVTSVKYKFTV-FNPSF-----VYMDRYASSL--GAPLHNSCISALFLIF 366
QY 1127 -SCLLGLDLRSLNL--LSIVMLVDTVGFMALWDISYNAVSLINLVSAGMSVERV 1182
Db 367 FRAFLV-----DSLNVWITLTVSVFVGFMILWKVELDCISVLCLYGINITYDNC 422
QY 1183 SHITRSFAIS---TKPTWLERAKEATISMSGAVFAGVAMTN-----LPGILVLGLAKAQL 1234
Db 423 APMLSFTVLGKDFTRTKWKNALV-----HGVALQSYLCYIVGLIPLAAVPSNL 473
QY 1235 IQIFFRMLNLLTLGLLHGLVFLPVLSYVGPDPNPALAEQKRAE 1282
Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFLPSS-----KKGRKEK 513

RESULT 24
US-11-360-355-141365
; Sequence 141365, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
```

```

; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 141365
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Heterodera glycines
;
US-11-360-355-141365

Query Match      3.2%; Score 222.5; DB 7; Length 422;
Best Local Similarity 25.1%; Pred. No. 1.9e-09;
Matches 78; Conservative 61; Mismatches 117; Indels 55; Gaps 13;

QY 540 LSCNADYGAPVFPFLAIGYK----GKDYSBAEALIMTFSLNNYPAGD--PRLAQAK--L 591
Db 1 MNCPED-----LIVGIRRDAGGVQIRHAEAFQTVFLVAS--AGDVYQRLIRSKNHR 50
QY 592 WEEAFLEEMRAPQRMAGMFQVTTAERSLEDEINRTTAEDL----PIFATS----- 639
Db 51 MEKFGVSTQAHAGAITAAWQNRFT--KSIYDHTLNKPAQGIRVVHPLASTSIQDMLBQF 108
QY 640 ----YIVIFL-YISIALGYSYSSWSRV-----WVDSKATILGLGGVAVVLGAVMAAMGFF 687
Db 109 SEQFFVIFIGVLMII--YAGMSQVHWQWFSVKSSCLLAIIGVLITLASVAGLGUS 166
QY 688 SYLGIRSSLVILQVVPFLVLSVGADNIPIFVLEYQRLPRRPGCEPREVHIGRALGRVAPSM 747
Db 167 TAMNHFNAATQIVPFLTLGLIGDDMPELLHNYNDVLE---AVRQKEVAVLLKETGMSV 223
QY 748 LLSLSSEAICFPLGALTMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDSKROEASRL 807
Db 224 LITSINNIAFETGCIIPALRSFCGQVAILTSLNVLCCIILFPAPFALDLRRKKAHR 283
QY 808 DV-----CCC 812
Db 284 DMSFAALFCCC 294

RESULT 25
US-10-461-673-16889
; Sequence 16889, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Ausudi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yuning
; APPLICANT: Chen, Rui-Hong
```



```

; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 16889
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-461-673-16889

Query Match      3.2%; Score 219; DB 6; Length 1137;
Best Local Similarity 20.8%; Pred. No. 1.7e-08;
Matches 156; Conservative 98; Mismatches 304; Indels 192; Gaps 28;

QY 621 LEDEINRTABDL-----PIFATSYIVIFLYISALGSYSWSRVMVD 663
DB 198 LADNTSVTGMGLKQELLRHLFLVQDTVYPLA-----LVAIFPGMALYLSLEFLTMV- 252
QY 664 SKATGLGG--VAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSVGADNIRIFVLE 720
DB 253 ----LLGVLSGLLVAFPLYQVAFRMAYFPFVNLAALL-----LSSVCANHTLIF-PD 300
QY 721 YORLPRR--PGBPREVHIGALGRVAPSMLLCSLSBAICFFLGALTPMPAVRTFAITSLG 778
DB 301 LWRLSKSLPSCGLAQVRGRTMHFGYLLLVSLGTTSAAFYASYLSRLPAVRCLALFMCT 360
QY 779 AVILDELLQMSAFVALLSDSKRQEARLSDVCCVKPQBLPPQGEG-----LLIGFPQ 833
DB 361 AVLVLHAL-----TLVWLPASAVLHERYLARGCARRAR---GRWEGSAPRRLIALHR 410
QY 834 KAYAPLLHWITRGVLLFLALFGVLSYSMCHISV-----GLDQELALP- 878
DB 411 LRLG---LRAAAGTSRLLPQRLPCGVTKFYIMICWFAALAAGAYTAGVSPRLRPT 467
QY 879 -----KDSYLLDYF-----LPLNRYFVGAPVYFVTTIGYN----- 909
DB 468 LPPPGQVFRPSHPFERFAEYRQLFLFEQLPQEGGHPVVLVWGVLVDGTGDLDPDS 527
QY 910 -----FS-----SEAGMAICSSAGCNPFSTQKIQTAYTEFPESYLAIPASSWVD 955
DB 528 NSSLRVDPAFSAAGPEAQWLLALCHRA--RQSPFDTLQEGW-----PTLCFVE 575
QY 956 DFIDWLTSPSCCLYISGNKDKFCBPSTVNSLNCNCKMSITWGSVRPSVEQPHKYLPH- 1014
DB 576 TLQRMWESPSCARL---GP--DLCCG-----HSDFPWA 603

1015 ---FLNDRNIRKCPKGLAAYSTSVNLTSDGQVLSRFRMAYHKPLKNSQDYTEALRAARE 1071
DB 604 PQFFLHCLKWMALQEGPDGTDGLGRFADAGHS--LAALVLQFQTNFRNSPDYNQTLFYNE 662
QY 1072 LAANITADLRKVRGTDPAEVEPPYITITNVFFYQYLITLPEGLFMLSCLIVPTFAVSCILL 1131
DB 663 VSHWLAARELGMAP---PGLRRGWFTSRLESLYLSQHSLSSTEPVVLGLALALAPAT--LLL 717
QY 1132 GL--DIRSGLLNLLSIWMILVDVTGFMALWDISVNAVSLNLSVAVGMSVEFYVSHITRSPA 1190
DB 718 GTWNVPLSLFSVAAGVAGTVLLTVGLLVLEWQNTARALFLSASVGLSVDFTVNYCISH 777
QY 1191 ISTKPTMLERAKEATISMGSVAVGAMTNLPGILVLGLAKAQLIQIFPRMLMLLTLGL 1250
DB 778 LCPHPDRLSRVAFSLRQTSCTATVGAALFAAGVLMPLATV-----LLYKRLG 825
QY 1251 LHLGLV-----FLPVILSYVGPVYN 1270
DB 826 IILMMVKVSCGFASFQSFQSLCCFPFPEKN 855

RESULT 26
US-10-461-673-10541
; Sequence 10541, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyang
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yunqing
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
```



```
Db 146 KF-----WLRDYEQFVEQDPEAR--IKFNLDKYEIKFFIXSAGLVFKLERS 191
QY 998 MGSVRSVEQFHKYLPWFLNDRNINIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLK 1057
Db 192 MRN-----ELAEFLQW-----PEF---QGFLSANGS-----ELVLSRFFF----- 223
QY 1058 NSQDYTEALR--AARELAANITADLRKVPCTDPAFEVFPVTITN---VFYEQYLTILPEG 1112
Db 224 TTASYGTLEBNSKREL---LLQWRVADVSRSP-VFLFTVQSDDAKFLDIPTMIPQT 279
QY 1113 -----LPMLSLC-----LVPTFAVSCLLGLDLRSGLNLLSIWV--ILVDT 1152
Db 280 IQSSACTLLCMFLVCLLEMTNPGALLVSNFAIFSTCIGV---FGIQLSLGTLDDPIFMSV 336
QY 1153 VGFMAWLDISYNAVSLINLVAAGMSVEFVSHITRSF-----AISTKPTWLERAKEAT 1205
Db 337 FGQSLGTLGDDPIFMSAIIAGMSFGSDVIPAIIAHFYKTSVGEKAGTSLDVEQRLMDCL 396
QY 1206 ISMSAVF-AGVA-MTNLPGLILVLGLAKAQLI-QIEFFRLNLLITLLGLLHGLVFLPVIL 1262
Db 397 ATIGFPVLEAGLSTLCTSSLPFVDLHMARVPARTVFAITVVAIGLLHGLLVIPVNF 456
QY 1263 S 1263
Db 457 S 457

RESULT 31
US-11-360-355-140206
; Sequence 140206, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 140206
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 66629; Strand=-; Position=1
; OTHER INFORMATION: -151,199-261,383-478,553-695,766-905,1296-1406
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAE71975.1; Match level="QueryCoverage
; OTHER INFORMATION: =91%, HitCoverage=25%, E-value=3e-82, Identity=68%"; Hit descrip
; OTHER INFORMATION: =Hypothetical protein CBG19049 [Caenorhabditis briggsae]
US-11-360-355-140206

Query Match 2.0%; Score 135.5; DB 7; Length 252;
Best Local Similarity 24.2%; Pred. No. 0.01;
Matches 45; Conservative 48; Mismatches 56; Indels 37; Gaps 10;

QY 1099 NVFYEQYL-----TILPEG---LFMLSCLV-----PTFAVSCLLGLDLRSGLNLLS 1144
Db 66 NFFSQMLEKNTTILUSSGTAILAMIFICVLFIADSSIVFWFSFMLSMDI--GVCGYL- 122
QY 1145 IVMLVDTVGFMAWLDISYNAVSLINLVAAGMSVEFVSHITRSFPAISTKPTWLERAKE- 1203
Db 123 -VSMDIGVCGYLSLWSDLDPTTVNVLMSIGLCIDFATHGYR-----TYSRCRDP 174
QY 1204 ---ATISMS-----AVFAGVAVTNLPGLILVLGLAKAQLIQIFPRLNLLITLLGLLHGLV 1256
```

```
Db 175 DQISDSLGAGWPVVOAGV--STFLGIIVMLLVPSHVVRM--AQTMLVAVATGLFHGLF 230
QY 1257 FLPVIL 1262
Db 231 LLIPIII 236

RESULT 32
US-10-724-972B-7050
; Sequence 7050, Application US/10724972B
; GENERAL INFORMATION:
; APPLICANT: DOUCETTE-STAMM, LYNN
; APPLICANT: BUSH, DAVID
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 47040.0007US
; CURRENT APPLICATION NUMBER: US/10/724,972B
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; NUMBER OF SEQ ID NOS: 7546
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 7050
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-724-972B-7050

Query Match 1.8%; Score 126; DB 6; Length 808;
Best Local Similarity 19.9%; Pred. No. 0.35;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 69 LNMHPHGMFNLYVDGLGLLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 128
QY 874 ELALP-KDSYLDY-----FLFNRYFEVGPVY-----FVTLTYNFSSEAGM 916
Db 129 MGVVLSNDPIILYLFWELTSFSPFLISFWREKASIYGAQKSLITVLG-GLSMLGGI 187
QY 917 NAICSSAGCNFSTQKIQVATFPEQSYL-----AIPASSWVDDFDW 960
Db 188 --ILLSLATDTFTSQAMISKASDIQSPFFILWMLFMIGAFKTSQAQVFFVWLPDAMEA 245
QY 961 LTPSSC-----CRLYISGPNKDKFCS-----TVNSLNLKNCMSITMGSVRPSVE 1006
Db 246 PTPVSAVYLSHATWVKAGLYLIIARITPIFAISEGWVWTIT-----LVGLITLFWASLNATKQ 301
QY 1007 QFHKYLPWFLNDRPNITKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEAL 1066
Db 302 HDLK-----GILAFSTVQLGMINSLMGLIGAVSYHYQGANSQLYAGF 344
QY 1067 RAARELAANI-----TADLRKVPCTDPAFEVFPVTITNVPFYEQYLT 1107
Db 345 VAAIFHLINHATFKGALFMITGGIDHSHGTGRDVKLGG---LLTIMPISFT-----LT 394
QY 1108 ILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDTVGFMAWLDISYNAV 1167
Db 395 VIT-----TSLMAGVPPF--NGFLSKFKFLESMLNVTHLMSLNTLIG-----ILLPIIA 442
QY 1168 LINLVAAGMSVEFVSHITRSFPAISTKPTWL--BRAKEATISMGSAVPAVMTNLPGLV 1226
Db 443 IIGSIPTFVYSIKFILHI---PFGSYKPEALPQAHESILM---LISPILLTSL--VIV 494
QY 1227 LGLAKAQLIQIFPRLNLLITLIG-----LLHGLVFLPVILSYVG 1266
Db 495 FGLFPSSILTSIITEPASVAVSQTSNTIABFHLPHGIT--PAFLSTIG 539
```

RESULT 33
US-11-360-355-149589
; Sequence 149589, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 149589
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 76012; Strand=-; Position=1
; OTHER INFORMATION: -51,383-466,817-1002,1773-1886,1992-2069
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAB6666.1; Match level="QueryCoverage
; OTHER INFORMATION: =89%, HitCoverage=16%, E-value=2e-48, Identity=63%; Hit descrip
; OTHER INFORMATION: =Hypothetical protein CBG12004 [Caenorhabditis briggsae]
US-11-360-355-149589

Query Match 1.8%; Score 123.5; DB 7; Length 171;
Best Local Similarity 26.3%; Pred. No. 0.053;
Matches 41; Conservative 38; Mismatches 62; Indels 15; Gaps 7;
QY 614 TPTAERSLEDEINRTTAEDLPFATSYIVIFLY-----ISLALGS-YSSWSRVVDSKAT 667
DB 12 TLYSQTLAEELKRNADTLVPFVFAFFILMFSLCNMSTCRGNYIDWTL-----SKPV 67
QY 668 IGLGVAVVGLVAMAMGPFSLGIRSSVLQVFPFLVLSVGADNIFIFVLEYQRLPRR 727
DB 68 LAILGVANGMG-IVTAIGLLNCMSVPYN-DIVGVMPFLVAVGVDMFLMIAAARVTRN- 124
QY 728 PGPPEVHIGRALGRVAPSMML-CSLSEACFFLGA 762
DB 125 -ASPVCRMRGEAMADAISWITITIKTLITISFGGA 159

RESULT 34
US-10-536-606-20
; Sequence 20, Application US/10536606
; GENERAL INFORMATION:
; APPLICANT: Cosson et al.
; TITLE OF INVENTION: VIRULENCE GENES, PROTEINS, AND THEIR USE
; FILE REFERENCE: 25421-502NATL
; CURRENT APPLICATION NUMBER: US/10/536,606
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: PCT/CH2003/00836
; PRIOR FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 10/324,967
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-536-606-20
Query Match 1.8%; Score 121.5; DB 6; Length 1043;

Best Local Similarity 17.7%; Pred. No. 1.2;
Matches 236; Conservative 166; Mismatches 398; Indels 531; Gaps 58;
QY 134 SLPIINVTR---VAQLGAGOLPAV---VAYEAPYQHSAPAEQSDVSCSRVRVPAATAAVG 186
DB 16 ALFISLAGLLVLSKLPVQAQYPNVAPPQITITATYP-----GASAKVLVD 59
QY 187 TMCVGYGSALCNAQRWLNFGQ-DTGNGLAPLDITFHLLEPG----- 226
DB 60 SVTSVLEESLNGAKGLLYPESTNNSGTABIVVTP---EPGTDPLDAQVDVQNRLLKAE 116
QY 227 ---QAV-GSGIQ-----PLNEGVARCNESQGDVATCSCQDCAASCAPAIAR 268
DB 117 RMPQAVLTQGLQVEQTSAGFLLIYALSKEGAQRSDTTALGDYAAARNINNELRLFGVGK 176
QY 269 PQALDSTFVLGQMPGSLVLIILCSVFAVVTILLGFRVAPARDKSKMDPKK---GTS 324
DB 177 LQFFSSE-----AAMRV-----WIDPQKLVGFGLS 201
QY 325 LSDKLSFSTHTLLGQFQGWGTWVASWPLTILVLSIPVVAL--AAGLVFTELTDDPVEL 382
DB 202 IDD---VSNAIRGQNVQ-----VPAGAFGSAPGSSAQELTATLAVK 239
QY 383 WSPNSQARSEKAFHDQHFPPFRTNQVILTPAPNSSRYDSLLGPKNFSGILDJDL 442
DB 240 GTLDDPQSGVYVLRANEDGSLVRLADVARLEBKESYNISSRLNGTPTVGGAIQLSPGA 299
QY 443 ELLE---LQERLHLQVMSPEAQRNISIQDIQYAPLNPDTNLSY-----DCINSLLQ- 492
DB 300 NAIQTATVLQKRLAELSAPFFE-----DMQYSV--PYDTSRFVDVAIEKVIHTLISA 349
QY 493 -----YFQNNRTLL--LTANQTLMGTSQVDWKDHPFLYCANAPLTFKDGFTALAL 540
DB 350 MYLVFLVMFLPLQNVRYTLIPSIIVPVCLGLTLMV-----YLLGFSVNMTMFGVNLAI 404
QY 541 SCADYGAIPFPFLAIGGYKDYSEARALIMFSLNNYPAGDPRLAQAOKLWEAPLEEM 600
DB 405 GILVD-----DAIVVENVERIMA--EEGISPA---EATVKAM 437
QY 601 RAFORMAGMPQVTTAERSLEDEINRTTAEDLPF--FATSYIVIFLYISLAL----- 651
DB 438 KQVSGAIVGITLV-----LSAVFLPLAFMAGSVGVIIYQFVSILVSLIFS 483
QY 652 -----GSYSWSR-----VAVDSKATLGLGGVAV 675
DB 484 GFALITFTFALCATLLKPIPEGHEKRGFGAFNGFARVTERYSILNSKLVARAGRFWL 543
QY 676 VLGAVMAAGPFSLGIRSSVLQVFPFLVLSVGADNIFIFVLEYQRLPRRPGPREVH 735
DB 544 VYAGLVAMLGYP-YLRLPEAFVPAEDLGMVVDV-----QLP--PGASR--- 584
QY 736 IGRALGRVAPSMMLCSLSEACFFLGAITPMPAVTEFA-----LTSGLAVILDLLQ 787
DB 585 -----VETDATGBELRFLKSRREAVSFLIS 611
QY 788 MSAPVALLSLDKRQASRLDVCCCKPQELPPQCGEGLLIGFPQKAVAPFLLHWITRG 847
DB 612 GFSP-----SQGDNAALAF-----PTFKDSEAG 636
QY 848 VLLFLFLALFGVLSYSMCHI SVGLDQRLALPKDSYLL----- 884
DB 637 -----AEQSAAEATAALNEHFALPDGTVMVAVSPPPINGLNSGGFALRLMDR 684
QY 885 -----DYFL-----FLNRYFE--VCAPVYFV-----TTLGYNSSSEAG 915
DB 685 SGVGREALLOARDTLLGHIQTNPKEFLYAMMEGLAEAPQLRLIDREKARALVSPFETISG 744
QY 916 -MNAICSSAGCNPSPFTOKIOYATFPPQSULAI PASSWVDDFIDWLPSSCCRLVISGP 974
DB 745 TLSAAGFGEVINDFTNAGRQQRVVIOAEQG-----NRMTPESVLELYV--- 787
QY 975 NKDFPCPSTVNSLNCNKMCSITM--GSVRPFSVEQFHKYLFWLNDNRPNIKCPKGLAAY 1032

Db 788 -----PNAAGNLVPLSAFVSVKWEGPV-----QLVRY-----NGYPSIRIVGDAAPGF 831
QY 1033 STSVNLTSGQVLASRFMAYHKPLKNSQDYTRALRAARELANITADLRKVPGTDPAFV 1092
Db 832 ST-----GEAMA-----EMERLASQLPAGI-----851
QY 1093 FPYTITNVFEQYLLTI-LPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLISVWIL-- 1149
Db 852 -GYEWTLGSLYQEKVSAGQATSLFALAILVV-----FULLVALYESWIPUSVWMLIVPI 903
QY 1150 --VDTVGFMALMDISYNAVSLINLVSAVGMS-----VEFVSHITRSPAISTKPTWLE- 1199
Db 904 GAIGAVLAVWVGMSNDVFKVGLITIIGLSAKNAIILIVEF-----AKELMEQG 952
QY 1200 -RAKEATISMGSAVFAVAMTNLP-----GILVL-----GLAKAQLIQIPFRLNLITLL 1249
Db 953 HSLRDAABEALRLFRPIITSMFAITLGVIPALASGAGASQRAITGTVIGMLSATPL 1012
QY 1250 GLIHLGLVLPV 1260
Db 1013 GVL-----FVPI 1019

RESULT 35

US-11-203-806A-11

; Sequence 11, Application US/11203806A

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Silverman, Joshua

; TITLE OF INVENTION: Nucleic and Amino Acid Sequences of

; FILE REFERENCE: Human Rf1 and Methods of Use Thereof

; FILE REFERENCE: 600-1-298N

; CURRENT APPLICATION NUMBER: US/11/203,806A

; PRIOR FILING DATE: 2005-08-15

; PRIOR FILING DATE: 2004-08-13

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 2426

; TYPE: PRT

; ORGANISM: Mus musculus

US-11-203-806A-11

Query Match 1.7%; Score 120.5; DB 6; Length 2426;
Best Local Similarity 19.9%; Pred. No. 5.2; Mismatches 106; Indels 341; Gaps 46;
Matches 184; Conservative 106; Mismatches 294; Indels 341; Gaps 46;

QY 11 LWALLRLAQSBPYTTIHQPGYCAFYDEGKPNELSGSLMTLSNVSCLSNTP-----62
Db 259 LWPLFVKLLGK---TLHRSG-----SFINSLLQLEELGFRSGTPMIKKIAP 302
QY 63 ARKITGDHLLLOKICPRLYTGPNTQACSAKQLVSLASLSITKALLTRCPACSDNPN 122
Db 303 AWKSLIDNPL-----NPDILCSAKRLKLLMQPLS-----S 333
QY 123 LCHNTCSNPQSLFINVTRVAOLGAGOLPAVVAEAFYQHSFAEQSDSCSRVVP-AAA 181
Db 334 IIVRTETLALTKEVWVWVLLMLRGP-QLPA-----NFEQVCVPLIQS 374
QY 182 TLAVGTCMGVYGSAALCNAQRWLNFGQDTGNGIAPLDITPHLEPGQAVGSGIQPLNEGVA 241
Db 375 TISVDSIPSPQG---NSSR-----GSASPGLSL-----LTPGH---KGASPYGSPRG 415
QY 242 RCNESQGGDVATCSQDCNCAACSPATARPQALDSTFYLGQMPGSLVLIILICSVFAVVTIL 301
Db 416 NUSSTNGG-----MAATPSI-----QLLG-----LEMLLHF 441
QY 302 LVGFRVAPARDSKMV---DPKKGTSLSKLSFS---THTLLQPFQGGWGTWVASWPLTIL 356
Db 442 LIGPEVLSPAKHQKIVLSLEPLEHPLISPSFSSKYAHTL-----ITAV 485
QY 357 VLSVIPVVAALAGLVFTELTDDVEL-WSAPNSQARSEKAFHQHGFPPFTNQVILTPAP 415

Db 486 HDSFVSVGKSDASDAVVSAINWELISLVKSVTEAGNRKEK-----SGSEVTL- 532
QY 416 NRSSRYRDSLLLGPKNF--SGILDLDLLELELEQER--LRHLQWSPBAQRNISLDICY 472
Db 533 -----LLKSLENIKSEVFPVSKTLVMEITVKGLPKVLGSPAYQ--VANWDI-- 579
QY 473 APLNPD-----NTSLYDCCINSLLQYFQNNRTLLLTANOTLMTGQTSQVDWKOHF 522
Db 580 --LNGTPALFLIQLIFNNLLCEGVED--EKFLNLETLV---GCVLSGPTSPLAFSDSV 632
QY 523 LYCAN-----APLT-----FKDGTALALSCHADYGAPVFPPL 554
Db 633 LTVINQNAKQLWKEHLWRMSMIVSPDVIHQTVNEVQGDALHNFSAIYGALTLPIN 692
QY 555 AIGGYKGDYGEAEALINTFSLNNYPAGDPRLAQAKLEWAEAFLEEMRAFORMMAGMFOVT 614
Db 693 HI-----FSAQTFTGTMK--ALLKTSELY---RAFTR---CASIV 726
QY 615 PTAER-----SLEDEINRTTAEDLPFPATSYIVIFLYISIALGSY-----654
Db 727 ATAENLCCBELSKIMCSLEDEV---LSDLFLDRISHIIIVMVDCIDFSPYNNKYQPK 783
QY 655 -----SSWSRVWVDSKATLGLGSAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVL 707
Db 784 IKSPQRSSDNR-----KKKEPLKGLASLPKLVKVIDTFTHTLSLK-----ETFSDTLL 832
QY 708 SVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVA-PSML-----LCSISEAICFFLG 761
Db 833 AIG--NSIISML-----SNVFGHISLPSMIREIFATFTPLALLYENS 873
QY 762 ALTPMPAVRTALTSGLAIVL-----DFLQMSAPVALLSLDSKQ-- 802
Db 874 KLDEAPKYVT--SLNNKLEKLGIEIVACLFQSYLGAYDSELEHLSPLLCVIFLHKNKQIR 932
QY 803 -----EASRLDVCCCVKPOELPP 820
Db 933 QQSALLWNATPAKATALVYPEELAP 957

RESULT 36

US-11-214-063A-2044

; Sequence 2044, Application US/11214063A

; GENERAL INFORMATION:

; APPLICANT: SUWA, MAKIKO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS

; FILE REFERENCE: 084335/166

; CURRENT APPLICATION NUMBER: US/11/214,063A

; CURRENT FILING DATE: 2005-08-30

; PRIOR FILING DATE: US/10/292,798

; PRIOR FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: 10/017,161

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: JP 2001-246789

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 2070

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2044

; LENGTH: 768

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-214-063A-2044

Query Match 1.7%; Score 118.5; DB 6; Length 768;
Best Local Similarity 25.9%; Pred. No. 1.3;
Matches 51; Conservative 36; Mismatches 87; Indels 23; Gaps 8;
QY 633 LPIFATSYIVIFLYISIALGSYSWSRVWVDSKATLGG-VAVVLGAVMAAMGFFSYLG 691
Db 3 LAFISSCIAALVYILTSCSVFLSFFGI-----ASIGLSCLVALFLYHVWFGI---QYLG 54

; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
; FILE REFERENCE: 13424 MOLECULES
; CURRENT APPLICATION NUMBER: US/11/312,958
; PRIOR FILING DATE: 2005-12-20
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-312-958-46

Query Match 1.7%; Score 116.5; DB 6; Length 619;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 107; Conservative 69; Mismatches 164; Indels 201; Gaps 26;
QY 536 TALALSCWADYCAPV-----PFLAIGGYKGDYSEAEALIMTFSLN---NYPAGDP 584
Db 77 SVLAGLCYAEFGARVPRSGSAYLSYVTVG-----ELWAFITGNLILSVIGTA 126
QY 585 RLAAQKLMEEAFLEEMRAFORRMAGMFQVTTAERSLEDEINRTTAEOLPIFATSYIVIF 644
Db 127 SVARA--WSSAF-----DNLIGN-HISKTLQCSIALHVPVLA-EYDFEALGLVLL 174
QY 645 LVISIALGSYSS-----WSRVMD----- 663
Db 175 LTGLLALGASESALVTKVFTGVNLLVLGFMISGFVKGDVHNWKLTEEDYELAMAEALNDT 234
QY 664 -SKATIGLGGV-----AVLGAVMAMGFFSYLG-----IRSLVIL 699
Db 235 YSLGLPGSGGFVFFGEGILRG---AATCFYAFVGFDCIATTGEBQNPQRSIPMGIVIS 291
QY 700 QVVPFLVLSVGDADNIFIFVLEYQRLPRRGEPRVHIGRALGRVAPSM-LIGSLSEACFCF 758
Db 292 LSVCFLAYFAVSALTLMWPYQLQPESLPEAFLYIGWAPARYVAVGSLCALSTSL-- 349
QY 759 FLGALTMPAV-----RTFA-----LTSGLAVILDFFLLQMSAFVA 793
Db 350 -LGSMPFPRVITYAMAEDGLLFRVLARIHTGTRTPIATVVSGLIAAFMAFLKLTDLVD 408
QY 794 LLSLDSKRQEASRLDVCCCV---KQGE----- 817
Db 409 LMSIGT-LLAYSLSVSCVILIRYQDPQETKTGEVELQBEAITTSEKLTLMGLFPPLNS 467
QY 818 LPPPGQGE-----GLLIGFFQKAYAPFLHWM-----ITRGVLLILFALFG--VSL 861
Db 468 ITPPLSGQIVVYVCSSLLAVLLTALCLVLAQMSVPLLSGLDLWTAVVLLLLIIGIIVI 527
QY 862 YSMCHISVGLDQEL-----ALPKDSYLLDYFLFLNR-----YFEVGAPVYFTYILG 907
Db 528 WRQOSSTPLPKVPALPLLPLMSIFVNIYLMQMONTAGTWARFGVMMLIGAIYF----G 583

QY 908 Y 908
Db 584 Y 584
RESULT 40
US-10-461-673-12167
; Sequence 12167, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yungqing
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: PTFL_genes Version 6.0
; SEQ ID NO 12167
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-673-12167

Query Match 1.7%; Score 116.5; DB 6; Length 619;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 107; Conservative 69; Mismatches 164; Indels 201; Gaps 26;
QY 536 TALALSCWADYCAPV-----PFLAIGGYKGDYSEAEALIMTFSLN---NYPAGDP 584
Db 77 SVLAGLCYAEFGARVPRSGSAYLSYVTVG-----ELWAFITGNLILSVIGTA 126


```
Db 453 -----GLTDRFTAGATDDGTTVLSS-----TGIQD-----RSINEQVVEQLRAI 491
QY 1082 KVP-----GTDPAFEV-----FPYITINVEYOYLFIPEGLFMLSICLVPTFAVSC 1128
Db 492 SVPEGEVVOIGGTPAMEIESIEALFEKLLWMAIYIVLATFILMALVFGSV-TLPAKAIIM 550
QY 1129 LLLGLDLRLSGLNLISIVMLVDTVGFMALWDISYNAVSLIN-----LVSANV--GMSVFEV 1182
Db 551 TILGAGATGILITLM-----FVDGVGASAL-----NFSPGPLMSPVLVIMAIYGLSTDY- 601
QY 1183 SHITSPALSTPTMLERAKEATISMGSAVPAGVAMTNLPGLVLGLAKAQIQLIQTFFPL 1242
Db 602 -----EVFLVSR-----MVEARDKGESTDDAIRYGTHT-----GSIITAAALIMIVVCGAGF 650
QY 1243 NLLIITLLGLLHGLV 1256
Db 651 SEIVVMKYIAFGMI 664

RESULT 44
US-11-360-355-152528
; Sequence 152528, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zifeng
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 152528
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_78951; Strand=+; Position=1
; OTHER INFORMATION: -147,625-754,987-1185,1232-1296
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_495662.1; Match level="QueryCoverage=79%, HitCoverage=13%, E-value=2e-46, Identity=57%"; Hit description: "patched family member (ptc-1) [Caenorhabditis elegans]"
; OTHER INFORMATION: emb|CAA86843.1| Hypothetical protein ZK675.1 [Caen
US-11-360-355-152528

Query Match 1.6%; Score 111; DB 7; Length 201;
Best Local Similarity 27.9%; Pred. No. 0.7;
Matches 34; Conservative 19; Mismatches 65; Indels 4; Gaps 2;

QY 592 WEEAFLEEM--RAFOR--MAGMFQVTTAERLSLEDEINRTTAEDLPATSYIVIFLYI 647
Db 53 WARAFTDSLYNHAFNRDRDIEQSAVVHPLASTSISDMLAEFCDFNVAITLAGYLLMLLYA 112
QY 648 SLALGSYSSWSVMVDSKATLGLGVAVVLGVAMAMGPFSSVLGIRSSVLIVQVPELV 707
Db 113 LYSQCRFDCCSLGVESAVGLALAGVFTVTMASIAGLGLATWLGINFNAATTQIVPFLAL 172
QY 708 SV 709
Db 173 GI 174

RESULT 45
US-60-781-953-21
```

```
; Sequence 21, Application US/60781953
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jeffrey D.
; TITLE OF INVENTION: In Vivo Induced Genes of Mycobacterium tuberculosis
; FILE REFERENCE: 05-1122
; CURRENT APPLICATION NUMBER: US/60/781,953
; CURRENT FILING DATE: 2006-03-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21
; LENGTH: 958
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-60-781-953-21

Query Match 1.6%; Score 111; DB 8; Length 958;
Best Local Similarity 19.7%; Pred. No. 7.5;
Matches 115; Conservative 60; Mismatches 195; Indels 214; Gaps 26;

QY 347 WVASWPLTILVSVIPVVALAAGLVFTBELTTPDVELMSAPNSQARSEKAFHQHGFPPFR 406
Db 28 WVA---LTIIVNVVAP-----QLQSVARTHSVALGPHDA 58
QY 407 TNQVILTAPNRSSYRYSLLGPKNFSGILDLILLELELERLHLQVWSPRAQRNIS 466
Db 59 PSLIAMKRIGKDFQOFDS-----DTTAMV-LLEGQEKL-----GDEAHR--- 96
QY 467 LQICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCA 526
Db 97 FYDVLVTKLSDQTH-----VOHIEN-----FMGD----- 121
QY 527 NAPLTFKDGDTALALSCMADYCAPVFPFLAIGGYKGYKDYSEARALIMTF-----SLNN 578
Db 122 ---PLT-----AAGSQSADGKAAYVQLNLTDGQGSQANESVAAQRIVDVSVPPIPKA 173
QY 579 YPAGDPRLAQAKLWEEAFLEENRAFQRMAGMFQVTTAERLSLEDEINRTTAEDLPAT 638
Db 174 YVTGPGPL-----GADRVVY-GDRSLH-----TITGISI--- 201
QY 639 SYVIFLYISLALGSYSSWSVMVDSKATLGLGVAVVLGVAMAMGPFSSVLGIRS--- 694
Db 202 AVIAIMLFI-----AYRSLAALI-MLLTVGL-----ELLAVRGIISTFAVNDLMGL 247
QY 695 SILVILQVVPFLVLSVGDNIFFIVLEYQRLPRRPGEPREVHTGRALGRVAPSMMLCSLSE 754
Db 248 STFTVNLVALTIAASTDYIIFLVGRYQE-ARATQNRREAYVTWEGGTAHVVLASGLTV 306
QY 755 AICFFLGLATPMPAVRTTALTSGLAVILDFLQMSAFVALLSLSKQREASRLDVCCCVK 814
Db 307 AGAMYCLGFTRLPYFNTLASPCAIGLVTVMLASLTAPAIIVASR----- 352
QY 815 PQELPPQGGEGLLGPFQKAVAPFLHWTGTVLL-----LFLALPGVSVLSYM 864
Db 353 -----FGLFDPKRAITKRRWRRTGTVVVRWPGVPVLAATLLIATLIGL----- 393
QY 865 CHISVGLDQELAPKDSYLLDYFLFNRYFEVGPVYFVTTLGY 908
Db 394 -----LALPK--YQTNV---NERY---IPSAAPSNI 419

RESULT 46
US-11-045-004-2027
; Sequence 2027, Application US/11045004
; GENERAL INFORMATION:
; APPLICANT: BUCHRIEGER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
```

:	APPLICANT:	GLASER, PHILIPPE
:	APPLICANT:	KUNST, FRANCK
:	APPLICANT:	COSSART, PASCALE
:	APPLICANT:	DANIELS, JUSTIN
:	APPLICANT:	GOEBEL, WERNER
:	APPLICANT:	KREFT, JURGEN
:	APPLICANT:	KUHN, MICHAEL
:	APPLICANT:	NG, EVA
:	APPLICANT:	VAZQUEZ-BOLAND, ANTONIO
:	APPLICANT:	DOMINGUEZ-BERNAL, GUSTAVO
:	APPLICANT:	GARRIDO-GARCIA, PATRICIA
:	APPLICANT:	TIERREZ-MARTINEZ, ALBERTO
:	APPLICANT:	AMEND, ALEXANDRA
:	APPLICANT:	CHAKRABORTY, TRINAD
:	APPLICANT:	DOMANN, EUGEN
:	APPLICANT:	HAIN, THORSTEN
:	APPLICANT:	BERCHE, PATRICK
:	APPLICANT:	CHARBIT, ALAIN
:	APPLICANT:	DURANT, LIONEL
:	APPLICANT:	PEREZ-DIAZ, JOSE-CLAUDIO
:	APPLICANT:	BAQUERO, FERNANDO
:	APPLICANT:	GARCIA DEL PORTILLO, FRANCISCO
:	APPLICANT:	GOMEZ-LOPEZ, NURIA
:	APPLICANT:	MADUENIO, ENCARNIA
:	APPLICANT:	PABLOS, BETRIZ DE
:	APPLICANT:	WEHLAND, JURGEN
:	APPLICANT:	KARST, UWE
:	APPLICANT:	ENTIAN, KARL-DIETER
:	APPLICANT:	HAUF, JORG
:	APPLICANT:	ROSE, MATTHIAS
:	APPLICANT:	VOSS, HANUT
:	FILE OF INVENTION:	LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
:	FILE REFERENCE:	05394.0018-02
:	CURRENT APPLICATION NUMBER:	US/11/045,004
:	CURRENT FILING DATE:	2005-01-28
:	PRIOR APPLICATION NUMBER:	10/637,657
:	PRIOR FILING DATE:	2003-08-11
:	PRIOR APPLICATION NUMBER:	10/257,023
:	PRIOR FILING DATE:	2002-10-08
:	PRIOR APPLICATION NUMBER:	PCT/FR01/01118
:	PRIOR FILING DATE:	2001-04-11
:	PRIOR APPLICATION NUMBER:	FR 00/04,629
:	PRIOR FILING DATE:	2000-04-11
:	NUMBER OF SEQ ID NOS:	2854
:	SOFTWARE:	PatentIn version 3.3
:	SEQ ID NO	2027
:	LENGTH:	494
:	TYPE:	PRT
:	ORGANISM:	Listeria monocytogenes
:	US-11-045-004-2027	

Query Match	1.6%	Score 110.5;	DB 6;	Length 494;
Best Local Similarity	21.9%;	Pred.No.3;		
Matches	53;	Conservative 41;	Mismatches 111;	Indels 37; Gaps 7;
Qy	1075	NITADLRKV-PGTDPAFEVFP----	VTITNVFVEQVLTLTLPGLFLMLSLCVLTFAVSC	1128
Db	131	NVTGDIKFLBEGTKTIIVDVTFPFWAGVSFLWLTIGEAVFHPLPGITWSIAKKKGTQTQLG	190	
Qy	1129	LLLGDLIRS-GLLNLLISVMILVDTVGFMAWDISYNNAVSLINLVSAVGMSVFVSHITR	1187	
Db	191	IVLGLTLVSPQLLNAYSVV--ETKAGDIPWDFGFGAQQVMI-----GYQAQVIPAIMA	241	
Qy	1188	SFAISTKPTWLERAKEATISMGSNVPAGVAMTNLPGLILVIG-----LAKAQLIQIFPRL	1242	
Db	242	GFLLAYLEILWRKFIPNAISMIFVPFFPALVPTVLAHVILGPICGWKITGDAISNNVYAGLT	301	
Qy	1243	NLLITLAGLHGLVFLPVILS-----YVGPDPNPALALQKRAEEAAVA	1286	
Db	302	GGLSWLFAALFGFYAFVUVVGLHHMTNATDQLMSQFGGTNPWFMTALSNIAGGSVALA	361	
Qy	1287	VM 1288		

Db 41 GAFFLLFLVLTLLVMPLLIAEFVIGRGGSDAVQAYKTLAPGTWKSLLGKGVVGCASIL 100
QY 338 GQFFQCGTGWASWPLTILVLSVIPVALAAGLV-----FTELTDTPVELWSAPN 387
Db 101 FGFYSVVGWMI-----ITYLIKT-----LAGGIAGENQASLLHDFQVTTANP---WISVG 147
QY 388 SQARSEKAHQHGFQFFRTNQVILTPARSSRYDSSLILGPKNFSGIIDLLELLEL 447
Db 148 ATIL-----FILLNVISRGVVGSGIERKSKFMPALF--ILFVLIIRSLTL 193
QY 448 QERLRHLQVMSPEAQRNISLQDICYAPLNPDNITSLYDCCINSILOVFQNNRTILLLTANQ 507
Db 194 -----PGAMEGVAFP-----LRPD-----FSHTAQTVLITLQ 222
QY 508 TL-----MGQTSQVMDKHFLYCANAPLTFKGTALALSCWADY-----GAPVFPFLAIG 558
Db 223 AFPSLSVGISVMVYSSYLNRSTSLP-----QSAISVSLMNVFVSLLAGLAIFP----- 271
QY 559 YKGKDYSEBALIMTSLNYPAGDPRLAQAKLWBEAFLEEMKRAFORRMAGFPQVTPAE 618
Db 272 -----AASFNITDAGEFL-----LFVILPSIFNQMPFGM----- 302
QY 619 RSLDEINRTAEDLPFATSIVIEPLYSIALGSSYSSMSRVMD-----SKATLGL 670
Db 303 -----LFFIIFLFLFAULT--SSFSMLEATVAPLMNAGVNRKQASLWM 345
QY 671 GGVAVVLGVAAMGFFSYLGIRSSVLV-----QVVPFLVLSVGADNIFIFVLEY 721
Db 346 -GLVILMAIPSAISP-----GWSVDQIFGLSIFDAADYLVSNILPVGA--LFIAIFVG 398
QY 722 QRLPRR---PGPREVHIIGRALGRVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSG 778
Db 399 YRLPRELLKKEFTTSSHFGKVPV-----WFLIKYAPIAIILVFLSATGL 446
QY 779 AVILDEL 785
Db 447 ----IDPL 450

RESULT 48
US-11-056-355B-73134
; Sequence 73134, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 73134
; LENGTH: 758
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(758)
; OTHER INFORMATION: Ceres Seq. ID no. 4958918
US-11-056-355B-73134

Query Match 1.6%; Score 109.5; DB 7; Length 758;
Best Local Similarity 20.5%; Pred. No. 6.9; Mismatches 211; Indels 261; Gaps 41;
Matches 145; Conservative 89; Mismatches 211; Indels 261; Gaps 41;
QY 670 LGGVAVVLGVAAMAG-----FFSYLGIRS-----SLVI--LQVVP-----FL 705
Db 20 LGIAFQTLGVVGDGMSPLVFSDFSKVPISREVVDVLGSLVITYTIAVIPLAKYVFPV 79
QY 706 VLVS-----GADNIFIFVLEYQRLPRRG-BPREVHIG-----RALG----- 741

Db 80 VLKANDNGEGTFALYSLICRYAKVNKLPNQPADBQISSFRKLTPPELERALGKEAL 139
QY 742 ---RVAPSMLLCSLSAICFFLGALTPMPAVRTFALTSGL-AVILDF---LQMSAFVAL 794
Db 140 ETGKYLKTLALLVLMTGSMIIGDGLTTPAMSVMSAMSGLQGEVKGFGTALVMSIVIL 199
QY 795 LSLDKRQASRLDVCCCKVKPQORLPPQGGEGLLGFFOKAYAPFLLHITRGVLLPL 854
Db 200 VALFSIQRP-----GTGK---VGFL---FAPVLALM-----PF 226
QY 855 ALFGVLSYMSCHISVCLDQELALPKDSYLLDYFLPLNR-----YFEVGAPVYFVT----- 904
Db 227 SLGAGIYNLLKYDFTVIRAL---NPFYIV---LFPNKNKQAMSAALGGCVLCITGAEM 280
QY 905 ---TLGVNFSSEAGMAICSSAGCNFPSTQKIQYATEPPEQS-----YLAIPAS-SWVDDF 957
Db 281 PADLGHFSVRSIQMAFTCVVPPCLLAYMGQAYLTKHPEASARIYDVSVPKSLFVPV 340
QY 958 IDMLTPSSCCRLYISGPNKDKFCPTSVNSLCKNCKMSITMGSVRSPVQFHK----- 1010
Db 341 IATLAAMTASQAMISA-----TFSCVKQAMAL---GCFPLKIIHTSKKRIQ 385
QY 1011 -YLP---WFLNDRPNIKCPKGLAAYSTSVNLTSDQVLASRPMAYHKLKNQSDTEAL 1066
Db 386 IYIPVINWFL---MIMC-----ILV-----VSIF 406
QY 1067 RAARELAANITADLRKRVGCTDPAFVFPVYITNVFYEQYLTILPEGLFMLSCLVPTFAV 1126
Db 407 RSTTHIAN-----AYGIAEV-----GVMVSTVLV----- 431
QY 1127 SCLLGLDLRSGLNLLSIVMILV-DTVGFMAL-WDISYNAVSLINLVSAGVMSVB--FV 1182
Db 432 -----TLVMLLIWQTNIFLALCPPLIFSGSVETIYLLAVLTKILRGWV 474
QY 1183 SHITRSPALSTKPTW-----LERAKEA--TISM-----GSAFAGVAMTNLPGLVLGL 1229
Db 475 PLVFATFTLVNMYIWNYSVLKYQSEVRERISDMFMRIGST-----LGTIRIPGI---GL 527
QY 1230 AKAQLIQ-----IFFRLNLLITLLGLLHGLVFLPVLTSYVGPDPNP 1271
Db 528 LYNELVQIGPSIFG---QFLILPALHSTIIP--VCIKVVPVVPV 568

RESULT 49
US-11-056-355B-73133
; Sequence 73133, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 73133
; LENGTH: 772
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(772)
; OTHER INFORMATION: Ceres Seq. ID no. 4958917
US-11-056-355B-73133

Query Match 1.6%; Score 109.5; DB 7; Length 772;
Best Local Similarity 20.5%; Pred. No. 7.1;
Matches 145; Conservative 89; Mismatches 211; Indels 261; Gaps 41;
QY 670 LGGVAVVLGVAAMAG-----FFSYLGIRS-----SLVI--LQVVP-----FL 705

```

Db      34 LGIAFOTLGVVYDGMGTSPLYVFSDFVSKPIRSEVDVLGALSIVYITIAVPLAKYVFPV 93
Qy      706 VLSV-----GADNIFIFVLEYQRLPRPG-EPREVIHG-----RALG----- 741
Db      94 VLKANDNGEGGTALYSLICRYAKVNKLPNQPADQISSFRKLKLPTELEALGIKEAL 153
Qy      742 ---RVAPSMLLCSLSSEACFFLGALTPMPAVRTALTSGL-AVILDF---LLQMSAFVAL 794
Db      154 ETKGYLKTLILLVLMGTSMIIGDGLTPAMSVMSAMSGLOGEVKGFGTNALVMSIVIL 213
Qy      795 LSLDSKRQESASRLDVCCCKVQBELPPPGQEGGLLGFQKAYAPFLLHWTIRGVVLLFL 854
Db      214 VALFSIQRF-----GTGK---VGFL---FAPVLALW-----PF 240
Qy      855 ALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLPLNR-----YFEGCAPVYFVT----- 904
Db      241 SLGAIGIYNLLKYDFTVIRAL---NPFYIV---LFFNKNKQAWMSALGCGVLCITGABAM 294
Qy      905 --TLGYNFSSSEAGMNAICSSAGCANNFSFTOKIQATEPPEQS---YLAIPAS-SWVDDF 957
Db      295 FADLGHFSVRSIQMAFTCVVFPCLLLAYMGQAAYLTKHPEASARIFYDSVPKSLFWPVFV 354
Qy      958 IDWLTSSCCRLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRPSVEQPHK----- 1010
Db      355 IATLAAMIASOAMISA-----TFSCVKQAMAL---GCFPRKLIHTSKRIGQ 399
Qy      1011 -YLP---WFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEAL 1066
Db      400 IYIPVINWFL---MIMC-----ILV-----VSIF 420
Qy      1067 RAARELAANITADLRKVPGTDPAPFVFPYITTNVFEQYLTILPGLFMLSCLVPTPAV 1126
Db      421 RSTTHIAN-----AYGIAEV-----GVMMVSTVLV----- 445
Qy      1127 SCLLGLDLRSGLNLLSIYVILV-DTVGFMAL-WDISYNAVSLINLVSAGVMSVE--FV 1182
Db      446 -----TLVMLLIQTNIFLALCFPLIFGSEVETIYLLAVLTKILEGGWV 488
Qy      1183 SHITRSEAFSTKPTW-----LERAKEA--TISM-----GSAVFAGVAMTNLPGILVLGL 1229
Db      489 PLVFATPFLTWYIWNYSVLYKQSEVRERISDMFRELGST-----LGTIRIPGI---GL 541
Qy      1230 AKAQILQ-----IFFRLNLLITLGLLHGLVFLPVLVSYVGPVNP 1271
Db      542 LYNELVQGISIFG---QFLLTLPALHSTIIF--VCIKYVPVPVVP 582

RESULT 50
US-11-056-355B-73132
; Sequence 73132, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 73132
; LENGTH: 827
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(827)
; OTHER INFORMATION: Ceres Seq. ID no. 4958916
US-11-056-355B-73132

Query Match 1.6%; Score 109.5; DB 7; Length 827;
Best Local Similarity 20.5%; Pred. No. 7.9;

```

```

Matches 145; Conservative 89; Mismatches 211; Indels 261; Gaps 41;
Qy      670 LGGVAVVLGAVMAAMG-----FFSYLGIRS-----SLVI--LQVVP-----FL 705
Db      89 LGIAFOTLGVVYDGMGTSPLYVFSDFVSKPIRSEVDVLGALSIVYITIAVPLAKYVFPV 148
Qy      706 VLSV-----GADNIFIFVLEYQRLPRPG-EPREVIHG-----RALG----- 741
Db      149 VLKANDNGEGGTALYSLICRYAKVNKLPNQPADQISSFRKLKLPTELEALGIKEAL 208
Qy      742 ---RVAPSMLLCSLSSEACFFLGALTPMPAVRTALTSGL-AVILDF---LLQMSAFVAL 794
Db      209 ETKGYLKTLILLVLMGTSMIIGDGLTPAMSVMSAMSGLOGEVKGFGTNALVMSIVIL 268
Qy      795 LSLDSKRQESASRLDVCCCKVQBELPPPGQEGGLLGFQKAYAPFLLHWTIRGVVLLFL 854
Db      269 VALFSIQRF-----GTGK---VGFL---FAPVLALW-----PF 295
Qy      855 ALFGVLSYMSCHISVGLDOELALPKDSYLLDYFLPLNR-----YFEGCAPVYFVT----- 904
Db      296 SLGAIGIYNLLKYDFTVIRAL---NPFYIV---LFFNKNKQAWMSALGCGVLCITGABAM 349
Qy      905 --TLGYNFSSSEAGMNAICSSAGCANNFSFTOKIQATEPPEQS---YLAIPAS-SWVDDF 957
Db      350 FADLGHFSVRSIQMAFTCVVFPCLLLAYMGQAAYLTKHPEASARIFYDSVPKSLFWPVFV 409
Qy      958 IDWLTSSCCRLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRPSVEQPHK----- 1010
Db      410 IATLAAMIASOAMISA-----TFSCVKQAMAL---GCFPRKLIHTSKRIGQ 454
Qy      1011 -YLP---WFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEAL 1066
Db      455 IYIPVINWFL---MIMC-----ILV-----VSIF 475
Qy      1067 RAARELAANITADLRKVPGTDPAPFVFPYITTNVFEQYLTILPGLFMLSCLVPTPAV 1126
Db      476 RSTTHIAN-----AYGIAEV-----GVMMVSTVLV----- 500
Qy      1127 SCLLGLDLRSGLNLLSIYVILV-DTVGFMAL-WDISYNAVSLINLVSAGVMSVE--FV 1182
Db      501 -----TLVMLLIQTNIFLALCFPLIFGSEVETIYLLAVLTKILEGGWV 543
Qy      1183 SHITRSEAFSTKPTW-----LERAKEA--TISM-----GSAVFAGVAMTNLPGILVLGL 1229
Db      544 PLVFATPFLTWYIWNYSVLYKQSEVRERISDMFRELGST-----LGTIRIPGI---GL 596
Qy      1230 AKAQILQ-----IFFRLNLLITLGLLHGLVFLPVLVSYVGPVNP 1271
Db      597 LYNELVQGISIFG---QFLLTLPALHSTIIF--VCIKYVPVPVVP 637

RESULT 51
US-60-752-355-33716
; Sequence 33716, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33716
; LENGTH: 3979
; TYPE: prt
; ORGANISM: Streptomyces nanchangensis
US-60-752-355-33716

Query Match 1.6%; Score 109.5; DB 8; Length 3979;
Best Local Similarity 20.8%; Pred. No. 86;
Matches 99; Conservative 57; Mismatches 144; Indels 177; Gaps 25;
Qy      376 TTDPVELWASPNQARSEKAFHDQ-----HFGPFRTNQVI-----LTAPNRSS 419

```

Db 2850 TTEAARAAPPKAKAVDITGFDRAAAGYHYGSPSYGGLQTVWROGEDLLADITLPTACT 2909
QY 420 --YRYSLLGLPKNFSGIILDL--LLELLEQLERLHLQVNSPEAQRNISIQDI---C 471
Db 2910 PDHTTDSLAIHP---ALLDAALHPLATADNPDG---EILWPTWSGVTLHATGATHV 2961
QY 472 YAPLNPNTSLYDCCINSLLQYFQNNRTLLLL-----T 504
Db 2962 RARITPOGDNDRYLTLD-----ATGQTVLTAGTIASTRPLDTRARLTRGPGDGLYQVRWT 3016
QY 505 ANQTLMGQTSQV--DW-----XDHELYCANAPLTPKDGCTALALSCNADYCAPVPEPL 554
Db 3017 AMPIPAGSATAVADDWAMLDGAGLDGGGLADAVAPLASYPDVVAALVAAMD--GTPV-PSV 3074
QY 555 AIGGYKGYKDYSAEALIMTFSSLNYPAGDPRLAQAKL-----WEEAFLEE----- 599
Db 3075 VLTGLPADGGDADVVEVLTTAREWLAEPRLAESRLVVVTHDAVAEDTSDPGDGDVD 3134
QY 600 -----MRAFO-----RRMAGM-----FQVTFARSL 621
Db 3135 PVAAGVWGLIRSAQENPGRTLLDLTRRDAGTAPDVVEVLRAAMDADWQVAVRGRAL 3194
QY 622 EDEINRTAEDLPIFATSVIVIFLYISLALGSYSSWSRVWDSKA----- 666
Db 3195 ---VPRLTAAD---AAGIV-----LPVGA-PAWQLVMADERAGTVDGLAPECEPEVL 3240
QY 667 -TLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVVP-----PLVLSVQAD 712
Db 3241 EPLAQGVQVIAVRA--AGVNF-----RDVMVTLGVPPDRRLGGEGAGWLDVAPD 3289

RESULT 52

US-60-752-355-39891
; Sequence 39891, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Adad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 39891
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Zea mays
US-60-752-355-39891

Query Match 1.6%; Score 108; DB 8; Length 474;
Best Local Similarity 19.2%; Pred. No. 4.5;
Matches 118; Conservative 76; Mismatches 210; Indels 212; Gaps 27;

QY 731 PREVH-----IGRALGRVAPSMMLCSLSAI-----CFFLGALTTPMPAVRTFALTSLGLAVI 781
Db 15 PRSEHTVAEAKRLLSLAGPLVASCTIQNVQLVSMFVGHGELPLAGA-SLASSLANV 73
QY 782 LDFLQMSAFVALLSLDSKROEASRLDVCCVKPQELPPPGQEGEGL-----LLGPFQK--- 834
Db 74 TGFSLLVGM-----ASALDTLC-----GOAFGARQYGLGLYKQRAM 110
QY 835 -----AYAPFLHLWITRGVLLLLFLALFGVSLYSNMCHISVGLDQSLALPKDSYLLDYFLF 889
Db 111 LVIALACVPIAAVWNAAGRILL-----LGQDRDIAEAGAYSRLWILS 154
QY 890 LMRYPEVGAPVYFVT-----LGVNFSSEAGMAIC-----SSAGCNFSPFTQKIQY 936
Db 155 LPVYVPLACHVRFLQTSIVPVMASSGATAGLHVLCWALVFKAGMSGKAALSGAISY 214
QY 937 ATEPPEQSVLAIPASSWDDFDIMLTPSSCCPLYISGPNKDF-----CPSTVNS 986
Db 215 SVNL-----AMLAL-----YVRLSSACKRTWTGFSFEAFRDLRLRFTELAVPSAM-- 258

QY 987 LNCLKNC---MSITGVSVRPSVEQPHKYLPMFLNDRPNI-KCPKGGLAAYSTSV-NLTSB 1041
Db 259 MVCLEWSPFELIALLSGLLPNPKLETSVLISICLNTGALLFMVYPYGLCTAISTRVSNELGA 318
QY 1042 GOVLASRFWAYHKPLKNSQDYTEALRAARELAANITADLRKVPDTPAPEVPPYTTNVP 1101
Db 319 GEPQAR-----LAARV----- 330
QY 1102 YEQYLITLPEGLFMLSCLIVPTFAVSCLLGLDLRSGLNLNLSIVMLIVDTVGFMALWDI 1161
Db 331 -----VNCIALSAGL-----LLGSTWILL-----RSFWGY 355
QY 1162 SY-NAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLERAK-EATISMSGAVPAGVA 1217
Db 356 MYSNEPEVVTYIARM-MPVLAISPFPTDGLHSLGSLGVLTCGRQKICARVNLGAYILAGIP 414
QY 1218 MTNLPGLVLGLAKAQILQIPPRNLNLTITLGLLHGLVFLPVLISYVCGPDVNPALALBQ 1277
Db 415 WA-----VLLAFVHLHNGMLGILGIVCGSLTKLVLLMWTIRIN-----WE 455
QY 1278 KEABEAAVAVMVASCP 1293
Db 456 KEATNAKETVPSSSLP 471

RESULT 53

US-10-724-972B-5410
; Sequence 5410, Application US/10724972B
; GENERAL INFORMATION:
; APPLICANT: DOUCETTE-STAMM, LYNN
; APPLICANT: BUSH, DAVID
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 47040.0007US
; CURRENT APPLICATION NUMBER: US/10/724,972B
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; NUMBER OF SEQ ID NOS: 7546
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 5410
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-724-972B-5410

Query Match 1.5%; Score 107; DB 6; Length 901;
Best Local Similarity 17.4%; Pred. No. 14;
Matches 102; Conservative 104; Mismatches 239; Indels 142; Gaps 24;

QY 260 AASCPAIAARPOALDSTF-----YLGQMPG-SLVLIILCSVFAVVTILLVGFVAPAR 311
Db 207 AAGMFIISAIIGLSSVGIALLTYIFDIPNFTLLAVMIGLAVGIDYSLFLFRKELK 266
QY 312 DKSRMVDPKGTSLSKLSFSTHTL-LQOFGQGTWVASWPLTILVLSVIPVALAAGL 370
Db 267 -----KKGVDVTVEAIAATAVGTAGSAVIFAGLTVMIATVAVCSLGVDFLAVMGFASAI 318
QY 371 -----VFTELTTDPVELMSAPNSQARSEKAFHQHFGFPFRTNOVILTAPNES-----SYR 421
Db 319 SVLFPAVLAALTLLPALI-----SIFHKSIKIKOKPTKSKDPKDH 358
QY 422 YDSLILGPKNFGIILDLLELLELLELRLHQLVNSPEAQRNISIQDICTYAPLNPDN 481
Db 359 WAKFIVGKPVIAVIVSL-IILILAAI-----PVSGHRLGIPDDSLKPTDSSEYK 406
QY 482 LYDCINSLQLQYFQNNRTLL-----TANOTLMGQTSQVDWKHFLYCANAPLTF 532
Db 407 AYKLIISDNFGCY-NGQIVMLVNTKDGSGKSTIERDLANNMRSDLEDIDNVDVTVSKAQLTD 465

```

533 KQGTALALSCMADYGAPVFPFFLAIGGYKGKDYSEAEALIMTFSLNNYTPAGDPRLAQAKLM 592
      |||      |||      |||      |||      |||      |||      |||      |||
466 NNNYAL-----FTIIPEKGPNSQSTENLV-YDLRDYH-----SQA-----499
      |||      |||      |||      |||      |||      |||      |||      |||
593 EEAFLBEMRAFQRMMAGMFQVTFPAERSLDEINRTTAEDLPIPATSYIVIFLXISIALG 652
      |||      |||      |||      |||      |||      |||      |||      |||
500 -----QEKYDYGTEISGGQSVINIDMSEKNNAI-----FVPGVIVVLAFFLIMIV- 545
      |||      |||      |||      |||      |||      |||      |||      |||
653 SYSWSGRVWVDSKATLGLGGVAVLGVAAAGFF-----SYLGRSSLVILQVW 702
      |||      |||      |||      |||      |||      |||      |||      |||
546 ----FRSILVPLKAVLGF-----ILSLMATLGTTLVIOHGFMGSLFGIENTGPLLAF 595
      |||      |||      |||      |||      |||      |||      |||      |||
703 PFLVLSVGA-----DNFIFVLVEYQRLPRPCEPREVHIG-RALGRVAPSMLLCSLSE 754
      |||      |||      |||      |||      |||      |||      |||      |||
596 P-VIIIGLLFGLADYELFMTVRVHEYSKTGNDHSIRVGIKESGFVIAAALMFPSV 653
      |||      |||      |||      |||      |||      |||      |||      |||
755 AICFFLGALTPMPAVRTALTSGSLAVILD-FLLQMSAFVALLSLDSK 800
      |||      |||      |||      |||      |||      |||      |||      |||
654 FIATF---VFQDQSAIKSMGIALGFCGLFDFAFVWRMTLIPALTKLFGK 697
      |||      |||      |||      |||      |||      |||      |||      |||

```

```

RESULT 54
US-10-567-867-1237
; Sequence 1237, Application US/10567867
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: US/10/567,867
; CURRENT FILING DATE: 2006-01-27
; PRIOR APPLICATION NUMBER: US 60/490,890
; PRIOR FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2786
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1237
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-567-867-1237

```

Query Match	1.5%;	Score 106.5;	DB 6;	Length 629;
Best Local Similarity	20.0%;	Pred. NO. 9.1;		
Matches 118;	Conservative 71;	Mismatches 165;	Indels 235;	Gaps 31
QY	522	FLYCANAPLTFKDGATALALSCWADYGAPV-----FPFLATGGYKGDYSEAEALIMT	573	
Db	70	FLIAALA-----SVLAGLCYGFEGARVPKTSAYLISYVTVG-----ELMAFITG	114	
QY	574	FSL--NNYPAGDPLRAQAALWBEAFLEENRAFORRMAGMF--OVTFTAERSLEDEINRTT	629	
Db	115	WNLLSYIIGTSSVARA--WSATPDELI-----GRPIGEFSRTHMTLNAQCVL-----	160	
QY	630	AEDLPFPATSYIVI-----FLYIS--LALSGYSWSRV	660	
Db	161	AENPDIFAVIIILITGLLTGLVKESAMVNKIFTCINVLVLGFIWVGFFVKGSVKWQLT	220	
QY	661	MVD-----SKATILUGGV-----NAVILGAVMAANGFSY-----	689	
Db	221	EEDFGNTSGRLCNDNTEKGKPGVGWMPFGFSGVLSSNAATCPYAFVGFDCIATTGEEVK	280	
QY	690	-----LGIRSSLTILQVPPFLVLSVGADNIFIVLEYGRLPRRPGCEPREV--HIGRALG	741	
Db	281	NPQKAIPGVIVASLLICFIAYF-----GVSAALTLMPYFCLDNNGFLPDAPKFKVWGGA	335	
QY	742	RVAPSM--LLCSLSEAI CFELGALTPMAV-----RT-----FALTS	776	
Db	336	KYAVAGVSLCALSAISL-----LGSMPFMRVITYAMAEDGLLKFPLANVNDRTKTPITLAS	392	

```

777 G-LAVILDFLLQMSAFVALLSLDSKROBASRLDVCCCV-----KPQ-----E 817
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
393 GAVAAVMAFLPDLKDLVLDLMSIGTLLAYSL---VAACVLVLRQPEQPNLVYQMASTSD 449
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
818 LPPPGQGE-----GLLGFFQKAVAPPLH 842
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 LDPADQNELASTNDSQGLFPEAEMPFLKTLSPKNMEPSKISGLIWNISLTAVLIT 509
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
843 W-----ITRGVLLFLFALFGVSLYSMMCHISVGL--DQELALPKOSYLLDYF--- 887
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
510 FCIVTVLGREALTKGALWAVFL-LAGSAL--LCAVVTGVIRWQPESKTKLSFKVPFLPVL 566
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
888 ---LFLNRY-----PBGAPVYFVTTILGYNF--SSAGHNA 918
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
567 PILSIFNVVYLMQLDQGTWVRFAVWMLGFIIFYGIGLWHSEASLDA 615

RESULT 55
US-60-752-355-16314
; Sequence 16314, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16314
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum ATCC 824
US-60-752-355-16314

```

Query Match	1.5%;	Score 106;	DB 8;	Length 435;	
Best Local Similarity	18.0%;	Pred. No. 5.7;			
Matches 119;	Conservative	84;	Mismatches 175;	Indels 282;	Gaps 31;
Qy	621	LEDEINRTT-----AEDLPFATSVIVFLVLSLALGS-----YSSMRVMVDSKATL	668		
Db	1	MEEKIKKETQIELSYGVEDPELLSRILLAPQHIFAFGGIIVPIIVASLSKLDATKST	60		
Qy	669	GLGGVAVVGLAV---MAAMGFFSYGLSRSLVLIVQVVPFLVSLVGADNIIFIVLEYORLP	725		
Db	61	ALLSAAILMAGVATLIQSKGVYS-VGARVACI-----MGDTITFV-----	99		
Qy	726	RRPGPREVHIGRALGRVAFSMLLSUSEBAICFLGALTMPAVRTFALTSGLAVIDFL	785		
Db	100	-----APAAV-VGKKFG-----LAGIFGAT--ILGAV-----IVVLSFF	131		
Qy	786	LMSAFVALLSLDSKRQEASRLDVCCCVKQELPPPGQGGELLGPFQKAYAPFLHWIT	845		
Db	132	IK-----:-----YIMKFPPI-----VT	144		
Qy	846	RGVLLLLFLALFGVSL-----YSWCHISVGLDQELAKPLKDSYLLDYELFLNRY	893		
Db	145	GIVVCLIGLTLTPSIDWAAGSGTKNYGSLQNISIAL-----IVMVITLLINH	194		
Qy	894	---FEVGAPYFVTTLTYNFSBAGMNAICSSAGCINNFSQTQI-----QYATEFP	941		
Db	195	GRGLVSSASILIGMWGY-----IICITPLGMVNFSTISNLKIISIPSIQYGVKEN	245		
Qy	942	EQSVLA-IPASSWVDIDWLTFPSSCCRLYISGPKDKFCPSIVNSLNCIKMCSITWGS	1000		
Db	246	ISALLPPIPAY-----FVSVI-----STVGNLRARNEISGI-----	276		
Qy	1001	VRPSVEQHFHYLPWFNLDRPNIKCPKGGLAAYSTSVNLTSD---GOVLASRFMAVHKPLKN	1058		
Db	277	-----KDDSKISKGVL-----SDGGISILAHGFGA-----	301		
Qy	1059	SDQYTEALRAARELAANIATADLRKVPQTDPAFEVFPFYITTNFVFEQYLTILPEGLFMLS	1118		
Db	302	-----MPNTSPSONVGLIPLTKV-ASRVYTLV-AGIILMVL	335		

QY 1119 CLVPTFAVSCLLGLDLRSGLNL-----SIVMI-LVDTVGFMAWMDISYNVSLI 1169
Db 336 GLVPRF-----SGIINIMQPVGVGVIMFGTVAAGIQTLSHVNLRNRL 383
QY 1170 NLVSAVGSVVFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGL 1229
Db 384 IIATSIGLGL-----GVTFRPEILSSLPESL-----KMFSSGISTGTIAALVLM 429

RESULT 56
US-11-027-399A-3216
; Sequence 3216, Application US/11027399A
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; APPLICANT: Zeng, Qiangdong
; APPLICANT: Opperman, Timothy
; APPLICANT: Houseweart, Chad Eric
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: 3687.1000-015
; CURRENT APPLICATION NUMBER: US/11/027,399A
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: US 10/640,833
; PRIOR FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: US 09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5324
; SEQ ID NO 3216
; LENGTH: 924
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-11-027-399A-3216

Query Match 1.5%; Score 106; DB 7; Length 924;
Best Local Similarity 18.6%; Pred. No. 18;
Matches 132; Conservative 99; Mismatches 228; Indels 252; Gaps 27;

QY 643 IFLYISLALGSYSSRVVDSKATLGLGVAVLGAVMAAMGFPSYLGIRSSVLQV 702
Db 407 VVLYAVAMVTITMTFR-FVDEERT-----HAGIFKALGYRSKDIK-- 448
QY 703 PFLVLSVAGADNIFIFVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSBAICPFLGA 762
Db 449 -FLLYGLVAGT-----VGTALGSILGHVLLASVSSV----- 479
QY 763 LTPMPAVRTFALTSGLV-----ILDFLLQMSAFVALLSLDSKRO---EASRL 807
Db 480 -----ITKGMVVGTEIQIFWTYVTSLLAFVLSLILASVLSAYLVAWRELHDEAAQL 528
QY 808 DVCCCVKQDELP-PPQCGEGLL-----GPFQKAYAPFLHWHITRGVVL-----LLFLALGV 859
Db 529 -----LLPKPPVKGAIIERIGFIWRRLS---FTHKVTARNIFRYKQRMMLTIFGV 577
QY 860 SLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPEVGPVYFVTTLTGYNFSSEAGMNAI 919
Db 578 A-----GSVAL-----LF-----AGLGTQ 591
QY 920 CSSAGCNFSTFQ-----KIQVATEPPEQSYLAIPASSWDDPIDWL 961
Db 592 SSVAGVPSKQFOIQOYQMLVSENPSATNQDKVELAEVLKQGEILAYQ----- 639
QY 962 TPSSCCRLYISGNKDKFCPSTVNSLNCILKNCMSITMSVRSVQFHKYLPWFLNDRPN 1021
Db 640 -----KIYSKTLDKDPKGAQLQNI-----TLMMIEKEDLTPTFILOHQQBELTLKQDIV 689

QY 1022 IKCPKGGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELANITADLR 1081
Db 690 ITAKLAQLAGVKGVTLEIEGK-----ELK 714
QY 1082 KVPCTDPAPFVPPYITNVFYEQYLITLPEGLPMLSLCLVPTFAVSCLLGLDLRSGLLN 1141
Db 715 VAAITENVYGVHFIY-MSQASYBQLYQLPQANTYL-VSLRDTSAT-ESQAGLLM 767
QY 1142 LLSIVMILVDTVGFMAWMD-----ISYNVSLINLVSAGVGSVVFVSHITRSFAISTKPTWL 1198
Db 768 NOSAVSSVVONASAIRLDFSDSIASSLNQMTILVIVSVLLAIVILYMLTNI-----NVA 820
QY 1199 ERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQLIQIFPFLNMLLITLGLLHGLVFL 1258
Db 821 ERIRE-----LSTIKVLGFHNE-VTLIYVRETIVLSLGVILGLIAG 862
QY 1259 PVLISVGVDPNPALEKRAEBAVAAVMAVASCNHPSPRSVSTADNIYNH 1309
Db 863 FYLHQLQIMISPATILFYPPQGVWEYVIVPAVAV-----SFILTLGLGFFVNY 909

RESULT 57
US-11-028-099A-3216
; Sequence 3216, Application US/11028099A
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; APPLICANT: Zeng, Qiangdong
; APPLICANT: Opperman, Timothy
; APPLICANT: Houseweart, Chad Eric
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: 3687.1000-019
; CURRENT APPLICATION NUMBER: US/11/028,099A
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: US 10/640,833
; PRIOR FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: US 09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5324
; SEQ ID NO 3216
; LENGTH: 924
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-11-028-099A-3216

Query Match 1.5%; Score 106; DB 7; Length 924;
Best Local Similarity 18.6%; Pred. No. 18;
Matches 132; Conservative 99; Mismatches 228; Indels 252; Gaps 27;

QY 643 IFLYISLALGSYSSRVVDSKATLGLGVAVLGAVMAAMGFPSYLGIRSSVLQV 702
Db 407 VVLYAVAMVTITMTFR-FVDEERT-----HAGIFKALGYRSKDIK-- 448
QY 703 PFLVLSVAGADNIFIFVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSBAICPFLGA 762
Db 449 -FLLYGLVAGT-----VGTALGSILGHVLLASVSSV----- 479
QY 763 LTPMPAVRTFALTSGLV-----ILDFLLQMSAFVALLSLDSKRO---EASRL 807
Db 480 -----ITKGMVVGTEIQIFWTYVTSLLAFVLSLILASVLSAYLVAWRELHDEAAQL 528
QY 808 DVCCCVKQDELP-PPQCGEGLL-----GPFQKAYAPFLHWHITRGVVL-----LLFLALGV 859
Db 529 -----LLPKPPVKGAIIERIGFIWRRLS---FTHKVTARNIFRYKQRMMLTIFGV 577
QY 860 SLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPEVGPVYFVTTLTGYNFSSEAGMNAI 919

Db	578	A-----GSVAL-----LF-----AGLGTQ	591
Qy	920	CSSAGCNFFSTQ-----KIYATEFPEQSILAIPASSWDDFDWL	961
Db	592	SSVAGVPSKQFOQQYOMLVSENPSATNQDKVELAEVLKGQEILAYQ-----	639
Qy	962	TPSCCRLIYISGPNKDFCEPSTVNSLNCLKNCHSIWTGVSVPSEGFHKYLFWFLNDRPN	1021
Db	640	-----KIYTLKDPKFGKGLONI-----TLMMEKEBDLTFFIHLQHRRQBELTKOGIV	689
Qy	1022	IKPCGGGLAAYSTSVNLTSQGCVLASRFMAYHKLXNSQDYTEALRAARELAANITADLR	1081
Db	690	ITAKLAQLAGVKVGQTLEIGK-----BLK	714
Qy	1082	KVPGTDPAFEVPPYTTNNVEYQVLTLPBGLPMLSCLVPTFAVSCLLGLDLRSLGN	1141
Db	715	VAAITENYGVHFY-MSQAQSYEQLYGQLPOANTYL-VSLRDTSATY-----IESQAGLLM	767
Qy	1142	LLSVMILTVDTVGFMALWD---ISYNAVSLINLVSAVGMSEVFVSHITSPASTKTPTWL	1198
Db	768	NQSAVSSVVQNASAIRLPDSIASLSLQNTWTILVIVSVLLAIVILNINI-----NVA	820
Qy	1199	ERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQIQIFFEFLMLLITLLGLHGLVFL	1258
Db	821	ERIRE-----LSTIKVLGFHNNE-VTLVIYRETIVLSLVGIVGLTAG	862
Qy	1259	PVILLSYVGPDVNPALALEQKRAEEAVANVWASCNHPSPRVSTADNIYNH	1309
Db	863	FYLHQFLQMISPATILFYDQGVGEVYVIPVAHV----SFILTLTGFFVNY	909
 RESULT 58 US-11-027-878A-3216 ; Sequence 3216, Application US/11027878A ; GENERAL INFORMATION: ; APPLICANT: Doucette-Stamm, Lynn ; APPLICANT: Bush, David ; APPLICANT: Zeng, Qiangdong ; APPLICANT: Opperman, Timothy ; APPLICANT: Houseweart, Chad Eric ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus ; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics ; FILE REFERENCE: 3687.1000-008 ; CURRENT APPLICATION NUMBER: US/11/027,878A ; CURRENT FILING DATE: 2004-12-30 ; PRIOR APPLICATION NUMBER: US 10/640,833 ; PRIOR FILING DATE: 2003-08-14 ; PRIOR APPLICATION NUMBER: US 09/583,110 ; PRIOR FILING DATE: 2000-05-26 ; PRIOR APPLICATION NUMBER: US 09/107,433 ; PRIOR FILING DATE: 1998-06-30 ; PRIOR APPLICATION NUMBER: US 60/085,131 ; PRIOR FILING DATE: 1998-05-12 ; PRIOR APPLICATION NUMBER: US 60/051,553 ; PRIOR FILING DATE: 1997-07-02 ; NUMBER OF SEQ ID NOS: 5324 ; SEQ ID NO 3216 ; LENGTH: 924 ; TYPE: PRM ; ORGANISM: Streptococcus pneumoniae US-11-027-878A-3216			
 Query Match 1.5%; Score 106; DB 7; Length 924; Best Local Similarity 18.6%; Pred. No.18; Matches 132; Conservative 99; Mismatches 228; Indels 252; Gaps 27;			
Qy	643	IFLYISIALCGSYSSSRVMDVSKATLGLGGVAVLGA VMAAGPFYSYLGRSSVLVLQV	702
Db	407	VVLVYAANA VYTFMTWR-FVDEERT-----HAGFKALGRVRSKDIIAK--	448
Qy	703	PFVLVSVGADNIPIFVLEYQRLEPRRGPREVHIHGALGRVAPSWMLCSISEACFTFLGA	762


```
; PRIOR APPLICATION NUMBER: 09/603,208
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 272
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-703-799B-272

Query Match      1.5%; Score 105.5; DB 6; Length 783;
Best Local Similarity 20.0%; Pred. No. 15;
Matches 145; Conservative 108; Mismatches 268; Indels 205; Gaps 34;

QY 628 TTAEDLPATSVIVIFLYISLALSGYSWSRWVDSKATGLGAVVILGAVMAAMGFF 687
Db 194 TSEIIGIGIAFVILFTF-----GSLIAAGLPLTAVIGVGIGALAVLAT----- 240

QY 688 SYLGIRSSVILQVDFVLSVGCADNIFVFL-----EYQRLPRPGEPREVHIGRALGR 742
Db 241 AFTDLNVTVPVLAVM--IGLAVGID-YALFILSRYAEYKMRP--ADAAGNAVGTAGSA 295

QY 743 V--APSMMLCSLSEAICFPLGALTMPAVRTALTSGLAVIDFLDQLQMSAFVALLSLDSK 800
Db 296 VVFAGATVIAALVALIADIIGFLTAM-----GISAAFTVFVAVIALTFTPALGVRFG 349

QY 801 RQEARLDCCCKVQPELPPCG-----EGLL-----LGFFQKAYAPFLH 842
Db 350 HAFKGI-----PGIGNPTPKQTWEQALNRRSKGRSNVKLVOK--APGL-- 392

QY 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVVF 902
Db 393 -----VVAVVVLGLGALTIPAM-----NLQLSLPSDS----- 419

QY 903 VTTLGVNFSSEAGMNAICSSACGNFSTQKIQTAYTEPEQSYLAIPASSWVDDPIDWLT 962
Db 420 TSNIDTQTSADLMAEGFGAGVN-----APPLVIVDTHVNDSTALQ 463

QY 963 PSSCRLYISGNKDKFCBPTVNSLNCNKMSITMGSVRPSVEQFHKYLPWFNDRENI 1022
Db 464 P-----LIEAQEPEGEF-----DREQAARFATYM--YVQTYSN 496

QY 1023 KCPGGLAAYSTSVNLTSDGQVLASRFMAYHKPL-KNSQDYTEALRAARELAANITADLR 1081
Db 497 NIDVKNAQIISVNDDFTA-AQILVT---PYTGADKETPELMHVLRAQE-----AQIE 545

QY 1082 KVPGETD---PAPEVFYITVVFYEQYLITLPEGLFMLSCLVPTTFVAVSCLLG-----L 1133
Db 546 DVTGELGTGTAVOLDIT-----EQLEDAMP-----VYLAVVGLAIFLILVFRSLIV 596

QY 1134 DLRSGLMLLSI-----VMILVDTVGFMAWD-----ISYNAVSLINLVSAVMSV---- 1179
Db 597 PLVAGLGLLSVGAAGFATVWQEGFGFVNTPGPLISFMPFIFLIGTVFGAMDYQVFL 656

QY 1180 -----EFVSHITRSPAISTKPTWLBRAKEATISMGs-----AVFAGVAMTNLP 1223
```

```
Db 657 VTRMRHYTHNGKQPGSKYTPVEQSVIEGTQGSRVVTAALIMIAVFVAFIDQPLPF 716
QY 1224 ILVLG--LAKAQLIQIPFFRLNLITLLGLL-HGLVFLPVLISVYVGPDVN-PALAEOKR 1279
Db 717 IKIFGALGAGVFFDAFFIRMLGVPASFMFLMGKATWMPKWLDRILPSLIDIEGTALEKEW 776
QY 1280 ABEAVA 1285
Db 777 BEKQAA 782

RESULT 62
US-60-751-420-2802
; Sequence 2802, Application US/60751420
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Van Berdewegh, Paul
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
; FILE REFERENCE: GENI-011/00US
; CURRENT APPLICATION NUMBER: US/60751,420
; CURRENT FILING DATE: 2005-12-19
; NUMBER OF SEQ ID NOS: 27266
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2802
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-751-420-2802

Query Match      1.5%; Score 105.5; DB 8; Length 1077;
Best Local Similarity 18.9%; Pred. No. 25;
Matches 167; Conservative 93; Mismatches 277; Indels 347; Gaps 40;

QY 607 MAGMFQVTFATRSLEDEINRTAEDLPATSVIVIFLYISLALSGYS-----SWSR- 659
Db 41 LAALLAVAWASGRELTSD-----PSFLTTLVLCALGGFSLLLGLASREQRQRWTRP 91

QY 660 ----VMVDKATLGLGVAVVILGAVMAAMGFFSYLGIRSSVILQVDFVLSVGCADNIF 715
Db 92 LSGLVWV---ALLALGHAFLTGTGVSAMQVSYF-----LF 125

QY 716 IFVLEYORLPRRPGEPREVHIGRALGR--VAPSMMLCSLSEAIC--PFLG-----ALT 764
Db 126 VIFTAYAWLP-----LGNRDAVAGLASSLSHLLVILGVLGPQPSRPALL 171

QY 765 PMPAVRTFALTSGLA-----VILDFLQMSAFVALLSLDSKR-----QEARLDVC 810
Db 172 POLAANAVFLCGNVAGVYHKALMERALRATFREALSSLSHRRRLDTEKKHQEHLLSIL 231

QY 811 CCVKPQELP-----PPGOG-----EGL-----LIGFFQKAYAPFL 840
Db 232 PAYLAREMKAIEIMARLQAGQSGRPESTNNFHSLYVKRHHQGVSVLYADIVGTR-----L 285

QY 841 LHMVTRGVLLFLALFG-----VSLYSMCH-----ISVG 870
Db 286 ASECSPKELVLMNELFGKFDQIAKEHECWRIKILGDCYCVSGCLPSLPLDPHAINCVRMG 345

QY 871 LDQ-----ELALPKDSY-LLDYFLFLNRYFEVGAPV 900
Db 346 LDMCRAIRKLRATGVTDINMRVGVHSGSVLCGVIGLQKQWQYDVMSHSDVTLANHMEAGV 405
```

```
QY 901 YFTVTLGVNFSBAGNNAICSSAGCNMFSTQKIQVATEFPFQSYLAIPASSWVDDFDIW 960
Db 406 GRVHTIGATALLAGAYAV-EDAGMEH-----RDPYLRELGBPTVLIDIPRAEEED-----455
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCCLKMCSITM-----GSRPSPVQEQHKYLP 1013
Db 456 -----EKGTAGGLLSLEGLKWRPSSLMTRYLESWGAAPPAHLSH-----496
QY 1014 WFLNDRP---NTKCPKGGGLAAVSTSVNL-----TSDG---QVLASRFMAYH 1053
Db 497 ---GDSVSTSTPLPEKTLASFSTQWSLDRSRTPRGLDDDELDTGDAKPFQVIEQ---LNSQ 551
QY 1054 KPLKNSQDYEARLAARELAANITADLRKVPGTDAPE-----VPPYTTITNVPEQYLT 1108
Db 552 KQWKOSKOPNPLTYLFXEKEBEKVRLSAI-----PAFKYEAETFLVFLSNFIQMLVTN 607
QY 1109 LPEGL-----FMLSCLLV-----PTFAVSCLLG-----LDRSGLLN-----1141
Db 608 RPPALAITYSITFLLELLIFVCFSEDLMRCVLKGPKMLHWLPALSGLVATRPGLRIALG 667
QY 1142 ----LLSIVMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTM 1197
Db 668 TATILLVPAMATSLTFFPPTSDCPQAPNVSSMISNL-----VLGLAKAQIQLIFFRLNLL 1245
QY 1198 LERAKEATISGMSAVFAGVAMTNLPGIL-----ELPGSLPLISVPYSMHCCTIGFLSCLSLFHMSPFLKLL 745
Db 708 -----LHGLVLP---VILSVVGP-DYNPALALBOK 1278
QY 1246 LLLLMLAASCSLFLSHAWLSCLVRLYLGLDSRPGVLEKPK 789
Db 746 LLLLMLAASCSLFLSHAWLSCLVRLYLGLDSRPGVLEKPK 789

RESULT 63
US-11-360-355-119639
; Sequence 119639, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 119639
; LENGTH: 308
; TYPE: prf
; ORGANISM: Heterodera glycines
; NAME/KEY: misc feature
; LOCATION: (121)..(121)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 46062; Strand=-; Position=1
; OTHER INFORMATION: -140,252-307,384-569,654-752,801-998,1285-1533
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit_ID=NP_494986.1; Match level="QueryCoverage=100%, HitCoverage=29%, E-value=5e-81, Identity=50%"; Hit description: Patched Related (ptr-6) [Caenorhabditis elegans] gb|AAA68302.1
; OTHER INFORMATION: Hypothetical protein C54A12.1 [Caenorhabditis elegans]
; FEATURE:
; OTHER INFORMATION: Essential gene: C elegans homolog=C54A12.1; Phenotype=Emb
US-11-360-355-119639
```

```
Query Match 1.5%; Score 105; DB 7; Length 308;
Best Local Similarity 18.9%; Pred. No. 4.1;
Matches 39; Conservative 29; Mismatches 66; Indels 72; Gaps 6;

QY 746 SMLLCSLEAICFPGLGALTTPMPAVTEALTSLGLAVILDELLQMSAFVALLSLDSK--RQE 803
Db 5 AIFPITSITDVVSFAAGTYSIDILAVRGFCMTAACMMFTPLYQITFFAALMVSISSRLQMOQ 64
QY 804 ASRLDVC--CC-----
Db 65 RNACNACLPCCLLAKDPPEFEBGKTELNNNNNNKNDKSSKAMDYEERDEKGFQKCKIKN 124
QY 816 QELPPPGQGEGLL-----GPFQKAYAPFLLHWTTRGVVLLFLA 855
Db 125 NTSAPISISSSSSALGASEFNLAMRQKPMGKFRNFYVNFYVDFLDMRTKFAVGTTFPV 184
QY 856 LFGVSLYSMCHISVGLDOR-LALPKD 880
Db 185 YLAISLWGLVSMQ-GLDYEKLLSSD 209

RESULT 64
US-11-056-355B-69790
; Sequence 69790, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69790
; LENGTH: 1080
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: (1)..(1080)
; OTHER INFORMATION: Ceres Seq. ID no. 1825823
US-11-056-355B-69790
```

```
Query Match 1.5%; Score 105; DB 7; Length 1080;
Best Local Similarity 19.7%; Pred. No. 27;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

QY 300 ILLVGRVA-PARDKSKMVDPKGTSLSDKLSFSTHTLLGQPFQGMG-----TWVASWP 352
Db 44 VOLIGFPVLFLFRPFSQTKNPK--PTEADFRKFSSTILGSVYIVTGLLVANSYSSVG 101
QY 353 LTLVLVSVVVALAAGLVFTTELITDTPVELWSAPNSQARSEKAFHDQHPGPPFRNQVIL 412
Db 102 LLLVSVSTFSLI-LASQLAFT-----APFSYFLNSQKFTPPVNSLFL 144
QY 413 TAPN-----RSSRYVDSLILGPNFSGIL-----DLDLLELLEL-QERLRHLQVWSP 459
Db 145 TISSALLVNTDSENTAKVRVYVIGICTIGASAGIGLLSLVQLILRKVKKQTPS- 203
QY 460 EAQRNISLQD-ICYAPLNPDTNLSYDCCINSLLQYFQNN--RTLLLTANQTLMGQTSQV 516
Db 204 -----TWTDLVAY-----QSLVASCV-VLIGLFASGEWKTLSSEMYKL-GKVPV 248
QY 517 -----DWKDFHLYCANAPLTFKDGITALALSCMADYGAVPFPLATGGYKKG-DYSEA 567
Db 249 MTLASIAISQVYTTIGVVG--LIFSSSVFSNSITA-VGLPIVPVAVIVFHDKNASKI 305
QY 568 EALIM-----TFSLNYY-----PAGDPRLAQAKLWEEAFLEEMERFQRMMAG 609
Db 306 FSIILAIWGFISFVYQHYLDEKLLKTSHTSPVGDPHLLPA-----EKG-----348
```

```
QY 610 MFQVTTAERSLE-----DEINRTTAEDLPATSYIVIFLYISLALGSSYSWRVMVD 663
Db : : : : :
349 --HTNIHSDONLEANLIDHEVVTSSSAVP-----QTENYKRWLRVSIY 391
QY 664 SKATLGLGGVAVVLGAVMAAGPFS--YLGIRSSVLQVVPFLVLSVGADNIPFIVLEYQ 722
Db : : : : :
392 VIFVLFCQPLATILGRLYYENGSTYV-----VTLQLIGFPVLV-----LFRP---FS 438
QY 723 RLPRRGEPREVHIGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTS-----GL 778
Db : : : : :
439 RIQPKSTDTNFSQSPSFTTLASVYLCTGLLSAVAYLSALA-FTAFPSYFLNSQKFTPL 497
QY 779 AVILDFLLQMSAFVALLSDSKQ--EASR-----LDVCCCKVQBELPPQGGILLGFPQ 833
Db : : : : :
498 IVSSLLLLTVSSALLVNTDSENSTNVRQVYVIGICTI-----GASAGIGLLSLTIQ 551
QY 834 KAYAPFLHWITRGVVLVLLFLALFGVSLYSMCHISVGL-----DOELALPKDS 881
Db : : : : :
552 MLFRKVTKHTSSAVT---DLAIY-QSLVASCVVILIGLPASGEWETLPSEMNYKLGKVS 607
QY 882 YLLD-----YPL-----FLNRYFEVGAPVY----- 901
Db : : : : :
608 YVLTLASAAISQVYTLGLVLIFESSVSFNSITAVGLPIVPAVAIVFHDRMDASKIF 667
QY 902 -----FVTLGYNFSSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db : : : : :
668 SIILACGLFSLFYQHYLDEKLTNTSHTSAVGDLHLFVEEGHTNIQSDONLEANLIDHEE 727
QY 943 QSYLAIPASS-----WV-----DDFIDWLTPSSCC--RLYISGPNKDKFCPSVNSLNCILK 991
Db : : : : :
728 TESFSVPQTKNCKRWLRVSIYAFVIFCQPLATVILGRLYYENGKSTYVTVTLQLIGFPV 787
QY 992 NCMSITMGSVR--PSVEQFHKYLPMFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db : : : : :
788 LILFRFESRIQPKSTDTNFSQSPSFTTLASVYLCT--GLLSAVAYLSAVGLLPLVST 845
QY 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVFPYIT 1098
Db : : : : :
846 FSLILASQLAFTAFPSYFLNSQKFTPLVNSILFLITVSSALL--VVNTDSE-----NTT 897
QY 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db : : : : :
898 NVSRQVYVIGICTIGASAGIGLVLSLIQLLFRKVTKHTSSAVLDLANYQSLVATCVVL 957
QY 1131 LGL-----DLRSLGLMLLSIVMLVDTVGFMALWDI--SYNAVSLI-----N 1170
Db : : : : :
958 IGLFASGEWRTLPSEMRNYKLGKVSILTLASAAIP---WQVYTVGCVLIFESSVSFNS 1014
QY 1171 LVSAVGMV 1179
Db : : : : :
1015 SITAVGLPI 1023
```

RESULT 65

```
US-11-056-355B-69789
; Sequence 69789, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69789
; LENGTH: 1097
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
```

```
; LOCATION: (1)..(1097)
; OTHER INFORMATION: Ceres Seq. ID no. 1825822
US-11-056-355B-69789
```

```
Query Match 1.5%; Score 105; DB 7; Length 1097;
Best Local Similarity 19.7%; Pred. No. 28;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;
```

```
QY 300 ILLVGRVA--PARDKSMVDPKGTSLSDKLSFSHTLLGQFPQGWG-----TWASWP 352
Db : : : : :
61 VOLIGFPVLFRFPFSQTKNEK--PTREADPKFSSFTILGSGVIYVTGLLVANSVMSVG 118
QY 353 LTILVLSVIPVVAALAGLVFTTELTTDPVELWSAPNSQARSEKAFHQDHGFFERTNQVIL 412
Db : : : : :
119 LLYLPVSTFSLI--LASQLAFT-----APFSYFLNSQKFTPLVNSLFL 161
QY 413 TAPN-----RSSRYVDSLGLGKPNFSGIL-----DLDLLELLEL--QERLRHLQVSWP 459
Db : : : : :
162 TISSALLVNTDSENTAKVSRVKTVIGICTIGASAGIGLLLSVLQILRKVLKKQTFP- 220
QY 460 EAQRNLSIQD--ICYAPLNPDNTSLYDCINSLLQYFQNN--RTLLLTANQTLMGQTSQV 516
Db : : : : :
221 ----TVDLVAI-----QSLVASCV-VLIGLFASGEWKTLTSEMENYKL-GKPYV 265
QY 517 -----DWKDFLYCANAPLTFKDTALALSCHADYGAFFFLAIGGVKKG-DYSEA 567
Db : : : : :
266 MTLASIAISMQVYTVGVG--LIFESSVSFNSITA-VGLPIVPAVAIVFHDRMNASKI 322
QY 568 BALIM-----TFSLNNY-----PAGDPLAQAKLWEEAFLEEMRAFQRRMAG 609
Db : : : : :
323 FSIILATWGFISFYQHYLDEKLTNTSHTSPVGDPHLLPA---EEG----- 365
QY 610 MFQVTTAERSLE-----DEINRTTAEDLPATSYIVIFLYISLALGSSYSWRVMVD 663
Db : : : : :
366 --HTNIHSDONLEANLIDHEVVTSSSAVP-----QTENYKRWLRVSIY 408
QY 664 SKATLGLGGVAVVLGAVMAAGPFS--YLGIRSSVLQVVPFLVLSVGADNIPFIVLEYQ 722
Db : : : : :
409 VIFVLFCQPLATILGRLYYENGKSTYV-----VTLQLIGFPVLV-----LFRP---FS 455
QY 723 RLPRRGEPREVHIGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTFALTS-----GL 778
Db : : : : :
456 RIQPKSTDTNFSQSPSFTTLASVYLCTGLLSAVAYLSALA-FTAFPSYFLNSQKFTPL 514
QY 779 AVILDFLLQMSAFVALLSDSKQ--EASR-----LDVCCCKVQBELPPQGGILLGFPQ 833
Db : : : : :
515 IVSSLLLLTVSSALLVNTDSENSTNVRQVYVIGICTI-----GASAGIGLLSLIQ 568
QY 834 KAYAPFLHWITRGVVLVLLFLALFGVSLYSMCHISVGL-----DOELALPKDS 881
Db : : : : :
569 MLFRKVTKHTSSAVT---DLAIY-QSLVASCVVILIGLPASGEWETLPSEMRNYKLGKVS 624
QY 882 YLLD-----YPL-----FLNRYFEVGAPVY----- 901
Db : : : : :
625 YVLTLASAAISQVYTLGLVLIFESSVSFNSITAVGLPIVPAVAIVFHDRMDASKIF 684
QY 902 -----FVTLGYNFSSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db : : : : :
685 SIILACGLFSLFYQHYLDEKLTNTSHTSAVGDLHLFVEEGHTNIQSDONLEANLIDHEE 744
QY 943 QSYLAIPASS-----WV-----DDFIDWLTPSSCC--RLYISGPNKDKFCPSVNSLNCILK 991
Db : : : : :
745 TESFSVPQTKNCKRWLRVSIYAFVIFCQPLATVILGRLYYENGKSTYVTVTLQLIGFPV 804
QY 992 NCMSITMGSVR--PSVEQFHKYLPMFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db : : : : :
805 LILFRFESRIQPKSTDTNFSQSPSFTTLASVYLCT--GLLSAVAYLSAVGLLPLVST 862
QY 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVFPYIT 1098
Db : : : : :
863 FSLILASQLAFTAFPSYFLNSQKFTPLVNSLFLITVSSALL--VVNTDSE-----NTT 914
QY 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
```

Db 915 NVSRVQVIGFICTIGASAGIGLVLSLQLLPRKVFTHKTSASVLDLANYQSLVATCVWL 974
Qy 1131 LGL-----DLRGLNLLSIVMLVDTVGFMALWDI-SYNVSLI-----N 1170
Db 975 IGLFASGEWRTLPSEMRNYKLGKVSYLTLASAAIF--WQVYTVGCVLIFESSSVFSN 1031
Qy 1171 LVSAVGMV 1179
Db 1032 SITAVGLPI 1040

RESULT 66
US-11-056-355B-69788
; Sequence 69788, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69788
; LENGTH: 1128
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1128)
; OTHER INFORMATION: Ceres Seq. ID no. 1825821
US-11-056-355B-69788

Query Match 1.5%; Score 105; DB 7; Length 1128;
Beat Local Similarity 19.7%; Pred. No. 29;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

Qy 300 ILLVGFRA-PARDKSKVDPKKTSLSDKLSFSTHLLGQFFQWG-----TWASWP 352
Db 92 VOLIGFPLFLFRFSSQTKNPK--PTEADFRKFSFTILGVSIVITGLLVANSYMSVVG 149
Qy 353 LTLVLSVIVVVALAGLVFTLTDPVELSNAPNSQARSEKAFHDQHFQPFRTNQVIL 412
Db 150 LYLFLVSTFSLI-LASQLAFT-----AFFSFYFNSQKFTPIVNSLELL 192
Qy 413 TAPN-----RSSRYDLSLLGPNFSGIL-----DLDLLELLEL-OERLRHLQWSP 459
Db 193 TISSALLVVNTDSENTAKVSRVKYVIGIITIGASAGIGLLLSVLQILRKVLKKQTF- 251
Qy 460 EAQRNLSQD-ICYAPLPNDTSLDYDCINSLLQYFQNN--RTLLLTANQTLMOQTSQV 516
Db 252 -----TVDLVAY-----QSLVASCV-VLIGLFASGEWKTLTSEMYNKL-GKVPYV 296
Qy 517 -----DWKHFLYCANAPLTKDGTALALSCMADYCAPVFPFLAIGGYKGK-DYGEA 567
Db 297 MTLASIAISQWYTTGVVG--LIFESSVFSNITA-VGLPIPVVAVIVFDKWNASKI 353
Qy 568 EALIM-----TFSLNNY-----PAGDPRLAQAKLWEBAFLMEARAFORMAG 609
Db 354 FSIILAIWGFISFVQHYLDKELKLTSTSPVGDPHLLPA---EEG----- 396
Qy 610 MQVFTFAERSLE-----DEINRTAEDLPFPATSYIVIFLYISLGSYSSWSRVMVD 663
Db 397 ---HTNIHSDONLEANLIDHEVVTSSSAVP-----QTEYKRWLRVSIY 439
Qy 664 SKATIGLGGVAVVLGAVMAAGFFS-YLGISSVLQVWPPLVLSVGADNIFIFVLVYQ 722
Db 440 VIFVLPQPLATILGRLYYENGNTYV-----VTLLQLIGFPVLV-----LFRP---FS 486
Qy 723 RLPRPGBPREVHIGALGRVAPSMLLCSLSAICFFLGTALTPMPAVRTFALTS-----GL 778

Db 487 RIROPKSTDTNFSOSPSTFTLASVYLTGLLVSAYAYLSALA-FTAFSYFLNSQKFTPL 545
Qy 779 AVILDFIQMSAFVALLSLDSKRO-EASR-----LDVCCCVKQQLPPQCGGLLIGFPQ 833
Db 546 IVSSLLLTIVSALLVNTDSENSTNSRVQVIGICTI-----GASAGIGLLLSLIQ 599
Qy 834 KAYAPFLHWITRGVVLVLLFLALFGVSYLSMCHISVGL-----DQELALPKDS 881
Db 600 MLFRKVFTHKTSASV--DLAIV-QSLVASCVLIGLFASGEWETLPSEMRNYKLGKVS 655
Qy 882 YLLD-----YFL-----FLNRYEVCAPVY----- 901
Db 656 YVLTLASAAISQWYTVTLGLVGLIFESSSVFSNITAVGLPIVPAVAIVFHDRMDASKIF 715
Qy 902 -----FVTLTGYNFSSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db 716 SILLAICGFLSFVQHYLDEKLTNTSHTSASVGDHLHPVEEGHTNTQSDNLEANLIDHEE 775
Qy 943 QSYLAIPASS-----WV-----DDFIDMLTPSSCC--RLYISGPNKDKFCPSTVNSLCLIK 991
Db 776 TESFSPQTKNCKRWLRVSIVAIFVIFQPLATVLGRLYYENGKSTVYVTLQLIGRPV 835
Qy 992 NCMSTMGSVR--PSVEQFHKVLPMFLNDRPNIKCPKGGLA-----AYSTSVM-----TS 1040
Db 836 LILFRFPRIRQPKSTDTNFSQSPSTFTLASVYLTCT--GLLVSAAYAYLSAVGLLYLPVST 893
Qy 1041 DQVVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPDPAPEVFPYTT 1098
Db 894 FSLILASQLATFAPFSYFLNSQKFTPLIVNSLFLTVSSALL--VVNTDSE-----NTT 945
Qy 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db 946 NVSRVQVYIGFICTIGASAGIGLVLSLQLLPRKVFTHKTSASVLDLANYQSLVATCVL 1005
Qy 1131 LGL-----DLRGLNLLSIVMLVDTVGFMALWDI-SYNVSLI-----N 1170
Db 1006 IGLFASGEWRTLPSEMRNYKLGKVSYLTLASAAIF--WQVYTVGCVLIFESSSVFSN 1062
Qy 1171 LVSAVGMV 1179
Db 1063 SITAVGLPI 1071

RESULT 67
US-11-360-355-129465
; Sequence 129465, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijiang
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: IN PLANTS AND COMPOSITIONS THEREOF
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 129465
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on cDNA: vcdna=SeqID_55888; Strand=-; Position=1
; OTHER INFORMATION: -91,165-245,660-825
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit_ID=NP_491005.2; Match level="QueryCovera

OTHER INFORMATION: =95%, HitCoverage=9%, E-value=7e-31, Identity=60%, Hit description
US-11-360-355-129465

Query Match 1.5%; Score 104; DB 7; Length 122;
Best Local Similarity 24.1%; Pred. No. 1.2;
Matches 27; Conservative 27; Mismatches 38; Indels 20; Gaps 3;

QY 698 ILQVVPFLVLSGADNIFVLEYORLP-----RRP-----GEPREHIG-----737

DB 10 VLVCTPFLVLAIGVDVAYLMMHAWQRITDQMRKPHAWQRITDQMRKPIGNSRCSKLP 69

QY 738 RALGRVAPSMILCSISEAICFPLGALTMPAVRTPALTSGLAVLDLFLQWS 789

DB 70 KVLVETGPAVLISALTNICADAVGFTGSPETILCVGNMAIAVDPIYQIT 121

RESULT 68

US-11-360-355-149551
Sequence 149551, Application US/11360355

GENERAL INFORMATION:

APPLICANT: Boukharov, Andrey

APPLICANT: Du, Zifeng

APPLICANT: Guo, Liang

APPLICANT: Kovalic, David

APPLICANT: Lu, Maolong

APPLICANT: McCarter, James

APPLICANT: Miller, Nancy

APPLICANT: Williams, Deryck

APPLICANT: Vaudin, Mark

APPLICANT: Wu, Wei

TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF

FILE REFERENCE: 38-21(53885)

CURRENT APPLICATION NUMBER: US/11/360,355

CURRENT FILING DATE: 2006-02-24

NUMBER OF SEQ ID NOS: 171306

SEQ ID NO 149551

LENGTH: 238

TYPE: PRT

ORGANISM: Heterodera glycines

FEATURE:

NAME/KEY: misc feature

LOCATION: (186)..(186)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (194)..(194)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 75974; Strand=-; Position=1
OTHER INFORMATION: -114,485-520,583-627,956-1094,1610-1739,2041-2109,2471-2644

OTHER INFORMATION: Homolog annotation: Hit ID=CAE60539.1; Match level="QueryCoverage

OTHER INFORMATION: =97%, HitCoverage=25%, E-value=2e-54, Identity=50%; Hit description

OTHER INFORMATION: =Hypothetical protein CBG04166 [Caenorhabditis briggsae]

US-11-360-355-149551

Query Match 1.5%; Score 104; DB 7; Length 238;

Best Local Similarity 21.6%; Pred. No. 3.3;

Matches 47; Conservative 44; Mismatches 87; Indels 40; Gaps 10;

QY 556 IGGYKGYSEALIMTFSLNNYPAGDPLRAQAKLWEEAFLEENRQRMRMAGHFWQTF 615

DB 27 IGGWKDDISDYE-----MSISNY-----FKDKYIGQ-----KYIGQHIVL 63

QY 616 TAERSLEDEINRTAEDLPFATSY-----IVIFLYISALGYSWSKVMVD-SKATLGL 670

DB 64 TISTSVQVEVDAGN-----IIRTFGGLAWMLCSLFSNLSAFFPHQSIYKLPVAI 118

QY 671 GGAVVLGVAMAMGFYSYLSIRSLVILQVVPFLVLSVGADNIFIVLEVQ---RLPRR 727

DB 119 FACLCPPFAGSGTALGILFPAGVRHS-SILAVTPFLILALGVDDAFLMTHSWQLASKRRH 177

QY 728 PGEPEVH--IGRALGR-VAPSMILCSISEAICFPLGA 762

DB 178 EGVANDANKLINQSNINDGTGPAILISTLITNIFADLVGA 215

RESULT 69

US-10-724-972B-4902

Sequence 4902, Application US/10724972B

GENERAL INFORMATION:

APPLICANT: DOUCETTE-STAMM, LYNN

APPLICANT: BUSH, DAVID

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 47040.0007US

CURRENT APPLICATION NUMBER: US/10/724,972B

CURRENT FILING DATE: 2003-12-01

PRIOR APPLICATION NUMBER: 09/134,001

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: 60/055,779

PRIOR FILING DATE: 1997-08-14

PRIOR APPLICATION NUMBER: 60/064,964

PRIOR FILING DATE: 1997-11-08

NUMBER OF SEQ ID NOS: 7546

SOFTWARE: PatentIn version 3.3

SEQ ID NO 4902

LENGTH: 510

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-10-724-972B-4902

Query Match 1.5%; Score 104; DB 6; Length 510;

Best Local Similarity 22.2%; Pred. No. 11;

Matches 50; Conservative 42; Mismatches 93; Indels 40; Gaps 7;

QY 985 NSLNLCKNCMSITMGSVRPSVEQFHKYLPLWFLNDRPNIKCPKGGLAAYSTSVNL-----1038

DB 308 NSLSEPNADSYGSG-----WFINTDHLVFTGTVDNFSQILLNIRKSY 353

QY 1039 -----TSDQCVLASRFMAYH--KPLKNSQDYTEALRAARELANITADLRKVGPTDPA 1089

DB 354 GIVVLANTNSNQVTR---LAEHLNTQIMNRRHYTTI-----EEKVNOTKDMQLIIST--L 403

QY 1090 FEVPYITITNVFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDLRLSGLNLLSIVML 1149

DB 404 ADIFWVIPSILVFSKILK-LREGHIFIRKCLRTSIMPSSIILGTFVAMMILFYLLPLIILG 462

QY 1150 VDTVGFMAWDISYNAVSLINLVSAGVMSVDFVSHITRSPAIATK 1194

DB 463 DATWGFVLSWLPKLSKYLVSIVLAITMLLVWLSLISITYRSDKK 507

RESULT 70

US-10-159-257C-191

Sequence 191, Application US/10159257C

GENERAL INFORMATION:

APPLICANT: SHEN, BEN

APPLICANT: LIU, BEN

APPLICANT: CHRISTENSEN, STEVEN D.

APPLICANT: STANDAGE, SCOTT

TITLE OF INVENTION: GENE CLUSTER FOR PRODUCTION OF THE ENEDIYNE ANTITUMOR

TITLE OF INVENTION: ANTIBIOTIC C-1027

FILE REFERENCE: 407T-896020US

CURRENT APPLICATION NUMBER: US/10/159,257C

CURRENT FILING DATE: 2002-05-31

PRIOR APPLICATION NUMBER: 09/478,188

PRIOR FILING DATE: 2000-01-05

PRIOR APPLICATION NUMBER: 60/115,434

PRIOR FILING DATE: 1999-01-06

NUMBER OF SEQ ID NOS: 225

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 191

LENGTH: 692


```
; TYPE: prt
; ORGANISM: Streptomyces globisporus
; FEATURE:
; OTHER INFORMATION: orf27
US-10-159-257C-191

Query Match      1.5%; Score 104; DB 6; Length 692;
Best Local Similarity 20.7%; Pred. No. 17;
Matches 65; Conservative 42; Mismatches 105; Indels 102; Gaps 11;

QY 503 LTANQTLMGQTSQVQWQKDFLYCANAPLTFKDGKTALALSC-----MAD 545
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 IAAQELAAEGKVE-----APLESEGQALWVVVPLISDADIVATTKKVRDVA 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 546 YGAPVFPFLAIGGYKGKDYSEABALIMTFSLNYPAGDPRLAQAKLWEAFLEMRAFOR 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 ANAPGVAIEVGG-----PAGSTTDA-----157

QY 606 RNAGQFQVTFTHAERSLEDEINTTAEEDLPFATSVIVIFLYISLALGSYSSRWVVDK 665
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 158 --AGAPE-----SLDSMLMVTG-----LVAILLLITVRSPIMLLLEPL--- 195

QY 666 ATLLGGVAVVLGAVMAAGPFSYLG-----RSSLVILQVVPFLVLSVGADNIFIFVLEY 721
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 -SVGFASULTQGTMLA-----KYAGLPVDPQSSGVLM-----VLVFGVGTDAVALLIARY 246

QY 722 QRLPRPGEPREHVHIGRALGRVAPSMMLCSLSEAICFFLGTALPMPAVRTFALTSLGLAVI 781
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 REELRR--QDRHVAMKTLARRSGPALLASAGTIAIGLCLVLADYNSRSRMLGVAIGVV 305

QY 782 LDFLLQMSAFVALL 795
   : : : :
Db 306 CALLAMVTILPALL 319

RESULT 71
US-11-056-355B-69565
; Sequence 69565, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69565
; LENGTH: 757
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(757)
; OTHER INFORMATION: Ceres Seq. ID no. 1818832
US-11-056-355B-69565

Query Match      1.5%; Score 104; DB 7; Length 757;
Best Local Similarity 20.6%; Pred. No. 19;
Matches 145; Conservative 91; Mismatches 209; Indels 260; Gaps 41;

QY 670 LGGVAVVLGAVMAAG-----FFSVLGIRS-----SLVI--LQVVP-----FL 705
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 LGIAQTLLGVVYVYVFSVFSKVPVIRSEVDVLGALSIVYTIATVPLAKYVVF 79

QY 706 VLSV-----GADNIFVLEYQRLPRPG--BPREVHIG-----RALG----- 741
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 VLKANDNEGNASAPMLVKYAKVKNLPQQAQDSQISFRUKLPTPELRAIGLKEALE 139

QY 742 --RVAPSMMLCSLSEAICFFLGTALPMPAVRTFALTSGI--AVILDF--LLQMSAFVALL 795
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```


Db 60 AKNPVAGPKDGSIFRELALFQDYGHLPSFKALVDFMABIRGNKVTFDPNHIIVLTAG 119
QY 416 NRSSVRY-----DSELLGPKNPSGILDLDL-----LELELQ-----ERLRH 453
Db 120 STSANQTLMFCLAEIGDAFLPTPYPG-FDRDLKRTGVEIIVPQCTSSNNPQTEAL 178
QY 454 LQVMSPEAQRNLSLODICYA-PLNPNTSLYDCCINSLLQYFQNNRTLLLLTANQTLMOQ 512
Db 179 KQAYQBAKRNLRVKGVLVTPNSPLGTTMSRSELNLLVDFIKDN--MHLISDEIYSTG 236
QY 513 T-----SQVDWKDHFYLC-ANAPLTFKDGCTALALSCWADYGAIPFPFLAIG 557
Db 237 VYNSPGFVSIMELKDRNLKD-----CGANV-----WDRVHIIVYSKDLGLGPFVRGAI- 287
QY 558 GYKGDYSEAEALI-----MTFSLNNYPAGDPRLAQAOKLWEEAFL-----BEMRA 602
Db 288 -----YSENEVVAAATWSSFGVLVSQTOY-----LLSAMLGDKFKTKYIENLKR 335
QY 603 FORM-----AGMF-----QVTFATRSLEDEI 625
Db 336 LKRRQRNLVSGLOKAGISCLTKNNAGLCWDMRHLHSNTPEAEMLWKI 387

RESULT 75

PCT-US06-00964-7153
; Sequence 7153, Application PC/TUS0600964
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/00964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 7153
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
PCT-US06-00964-7153

Query Match 1.5%; Score 103.5; DB 1; Length 618;
Best Local Similarity 19.4%; Pred. No. 16;
Matches 110; Conservative 67; Mismatches 195; Indels 195; Gaps 22;
QY 758 PFLGALTMPAVRTEALTSGLAVIDLFLQMSAFVALLSLDSKQBSRLDCCVCKP-- 815
Db 79 FHMSSLDAVALNVFAGTNTLATV-----VCAFASDLYLGRYATVAG-----CVSTFI 127
QY 816 -----QELPPPGQGGELLLGPFQKAYAPFLHWHITRGVLLFLA-----LFGVS 860
Db 128 GMVILTMTAGVPALHPPPCGEGRCIG-----ATRGQAVLGLAFAPFIVAGAG 174
QY 861 LYSMCHISVGLDQELALPKDSYLLDYFLPLNRYFVGAPVYFVTILGYNFSSEAGNNAIC 920
Db 175 GIRPCSLPFGADQ-----FDPRTESGRRGI- 199
QY 921 SSAGCNNSFTQKIQVATEPPEQSYLAIPASSWVDDFIDMLTPSSCCRLYISG----- 973
Db 200 -NSFPNWTYFTLTIIVACASSTAIIVTQSSVSWVWGLAIPAAIMLASCALFPAGAGLYVRV 258
QY 974 -PNKDKF-----CPSTVN-SLNCLKNCMSITMGSVRPSVBQF----- 1008
Db 259 REGSPFPAGVARVAVAAFRKRSAAAPSDADESLFTRHASGVV--SRLPYTDQFRFLDKA 316
QY 1009 -----HKYLPWFILNDRPNI---KCPKGGLAAYSTSVNLTSDDQVLASRPM 1051
Db 317 AVVVDKSEVGDGHPKPNWRLCSLQQVEEAKCILRVVPMWLTCTI-----V 362
QY 921 SSAGCNNSFTQKIQVATEPPEQSYLAIPASSWVDDFIDMLTPSSCCRLYISG----- 973
Db 200 -NSFPNWTYFTLTIIVACASSTAIIVTQSSVSWVWGLAIPAAIMLASCALFPAGAGLYVRV 258
QY 974 -PNKDKF-----CPSTVN-SLNCLKNCMSITMGSVRPSVBQF----- 1008
Db 259 REGSPFPAGVARVAVAAFRKRSAAAPSDADESLFTRHASGVV--SRLPYTDQFRFLDKA 316
QY 1009 -----HKYLPWFILNDRPNI---KCPKGGLAAYSTSVNLTSDDQVLASRPM 1051
Db 317 AVVVDKSEVGDGHPKPNWRLCSLQQVEEAKCILRVVPMWLTCTI-----V 362
QY 1052 YHKPLKNSQDYT--EALRAARELANITADLRKVPCTDPAPFVPPYTTINVPYEOY---- 1105
Db 363 YIVAFQNTTYVILQAQSDRHLGGGGGAGSFEVP--PGSFTVFPMALAVIPIYDRLV 420
QY 1106 -----LTILPEGLPMLSLCLVPTPAVSCLLGLDLRSGLNLLSIVMILVDVTGFMALW 1159

Db 421 VPWARRLTGREGGITPLQ-----RMGVGNMALSIVLAMLVAAMAEKRER 462
QY 1160 DISYNVS-----LINLVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATIS 1207
Db 463 DLAAAGSPNTGRVSRQSAPFWLPQALAGLSAFNQ-----VSQTEFYTFEPFESMRS 515
QY 1208 M-GSAVFAGVAMTN-LPGILVILGLAKA 1232
Db 516 VAGSVLFSGLALSSYLSGVLVAVERA 542

RESULT 76

US-11-330-403-7153
; Sequence 7153, Application US/11330403
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 7153
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-330-403-7153

Query Match 1.5%; Score 103.5; DB 6; Length 618;
Best Local Similarity 19.4%; Pred. No. 16;
Matches 110; Conservative 67; Mismatches 195; Indels 195; Gaps 22;
QY 758 PFLGALTMPAVRTEALTSGLAVIDLFLQMSAFVALLSLDSKQBSRLDCCVCKP-- 815
Db 79 FHMSSLDAVALNVFAGTNTLATV-----VCAFASDLYLGRYATVAG-----CVSTFI 127
QY 816 -----QELPPPGQGGELLLGPFQKAYAPFLHWHITRGVLLFLA-----LFGVS 860
Db 128 GMVILTMTAGVPALHPPPCGEGRCIG-----ATRGQAVLGLAFAPFIVAGAG 174
QY 861 LYSMCHISVGLDQELALPKDSYLLDYFLPLNRYFVGAPVYFVTILGYNFSSEAGNNAIC 920
Db 175 GIRPCSLPFGADQ-----FDPRTESGRRGI- 199
QY 921 SSAGCNNSFTQKIQVATEPPEQSYLAIPASSWVDDFIDMLTPSSCCRLYISG----- 973
Db 200 -NSFPNWTYFTLTIIVACASSTAIIVTQSSVSWVWGLAIPAAIMLASCALFPAGAGLYVRV 258
QY 974 -PNKDKF-----CPSTVN-SLNCLKNCMSITMGSVRPSVBQF----- 1008
Db 259 REGSPFPAGVARVAVAAFRKRSAAAPSDADESLFTRHASGVV--SRLPYTDQFRFLDKA 316
QY 1009 -----HKYLPWFILNDRPNI---KCPKGGLAAYSTSVNLTSDDQVLASRPM 1051
Db 317 AVVVDKSEVGDGHPKPNWRLCSLQQVEEAKCILRVVPMWLTCTI-----V 362
QY 1052 YHKPLKNSQDYT--EALRAARELANITADLRKVPCTDPAPFVPPYTTINVPYEOY---- 1105
Db 363 YIVAFQNTTYVILQAQSDRHLGGGGGAGSFEVP--PGSFTVFPMALAVIPIYDRLV 420
QY 1106 -----LTILPEGLPMLSLCLVPTPAVSCLLGLDLRSGLNLLSIVMILVDVTGFMALW 1159
Db 421 VPWARRLTGREGGITPLQ-----RMGVGNMALSIVLAMLVAAMAEKRER 462
QY 1160 DISYNVS-----LINLVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATIS 1207
Db 463 DLAAAGSPNTGRVSRQSAPFWLPQALAGLSAFNQ-----VSQTEFYTFEPFESMRS 515
QY 1208 M-GSAVFAGVAMTN-LPGILVILGLAKA 1232
Db 516 VAGSVLFSGLALSSYLSGVLVAVERA 542

RESULT 77

US-60-752-355-25099
; Sequence 25099, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25099
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Vitis vinifera
US-60-752-355-25099

Query Match 1.5%; Score 103.5; DB 8; Length 759;
Best Local Similarity 19.5%; Pred. No. 21;
Matches 167; Conservative 97; Mismatches 264; Indels 329; Gaps 41;

QY	543	MADYGPVPFPLAIGYKGDYSEAEALWTFSLNNYPAGDPRLAQAQLWEEAF-LEENR	601
Db	93	MGVFGAIIIFLFL--GSYKGFSTKSEPCYNTGSLCK-----PALANALFSTIAFLIGAMT	145
QY	602	AFQRRMAGPQVTFTAERSLEDEINRTTAEDLPFATSVIVLF-----	644
Db	146	SVFSGFLGKMIATYANA-----RTTLEARKGVGKAFIVAFRSGVNGVFLAASGLLV	197
QY	645	LYISLALGS--YSWSGRVMVDSKATLGLGVAVVLGAVNAAMGFFSYLGIRSSLVLOV	702
Db	198	LYVSIHLFSLYIGDDWEGLYESITGYGLG-----SSMALFRVG-----GSI	240
QY	703	PFLVLSVGADNIFIFVLEVORL-----PRPGEPRVHIGRALGRVA--PSMLLCSLSBA	755
Db	241	YTKAADVGAD-----LVGKVERNIPEDPRNPVAVIAD--NVGDNVGIAGMGSDLFGSYAEA	295
QY	756	IC--FFLGLALT-----PM-----PAVR	770
Db	296	SCAALFVASISSFGISHDYTAMSFPLIISSVGIVVCLGTTLFATDPFVKNVSETEPSLK	355
QY	771	-----TFALTSLAVILFLOMSAFVALLSLDSKROBASRLDVCCCKVPOELPPQGG	824
Db	356	ROLLISTILTAGIAVVSFALP-SEFTLFNFGSTKTKVNMHLFFCVSI-----GLW	406
QY	825	EGLLLGFF-----QKAYAP-----FLHWITRGVLLFLALFGV----	859
Db	407	AGLVIGYITEYTSNAYSVPQDVADSCRTGAATNVI FGLALGYKSVIIPFAIAIAIYVS	466
QY	860	-----SLYSCHISVGLDQELA--LPKDSYLLDYFLFLNRYFEVGAPVYFVITLGYNFSSE	913
Db	467	FSMAAMYGIAVAALGMLSTIATGLAIDAY-----GPI-----SDN	501
QY	914	AGHNAICSSAGCNNSFTOKIQYATEFPBQS-----YLAIPASSWD-----	955
Db	502	AG--GLAENAG-----MSHKIRQRTDALDAAGNTTAAIKGFAIGSAALVSLALFGAFVS	554
QY	956	-----DPIDLTPSSCRLYISGNPKDKFCPSTVNSLNCNKCMSTMGSVRPSVQPHKY	1011
Db	555	RAGIETVDVLT PKVFI GLVIGAWLPWFSAWMTKSVG-----SAAKMWEEVRRQFT-	607
QY	1012	LPWFLNDRNIIKCPKGGLAAYSTSVNLTSDDQVLASRFMAYHKPLKNSQDYTEALRAARE	1071
Db	608	IFPLMEGRAK-----PDYATCKVISTD-----	629
QY	1072	LAANITADLRKVPDGTDPAFVFPYITINVFYEQYLITLPEGLFMLSCLVPTFAVSCILL	1131
Db	630	-----ASLEM-----IPPGALVMLTPIAGT-----LF	653
QY	1132	GLDLRSGLL--NLSISVIMLVDTVGFMALWDISYNAVSLINLVSAVGMVSVEFVSHITRSF	1189
Db	654	GVETLAGVLGSLVSGVQVAISNTGGAWD--NAKKYIE--AGAS-----EH	697

QY	1190	AISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVILGLAKAQLIQIFFRFLNLLITLL	1249
Db	698	ARSLGPKGSDPHKAAVI--GDTIGDPLKDTSGPS-----LNLLIKLM	737
QY	1250	GLLHGLVLPVLVLSYVG	1266
Db	738	A-VESLVFAPFFAAHGG	753

RESULT 78
US-11-045-004-1583
; Sequence 1583, Application US/11045004
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: CHARBIT, ALAIN
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENIO, ENCARNNA
; APPLICANT: PABLOS, BETRIZ DE
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
; APPLICANT: HAUF, JORG
; APPLICANT: ROSE, MATTHIAS
; APPLICANT: VOSS, HANUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1583
; LENGTH: 1136
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-11-045-004-1583

```
Query Match 1.5%; Score 103.5; DB 6; Length 1136;
Best Local Similarity 19.1%; Pred. No. 39;
Matches 113; Conservative 84; Mismatches 190; Indels 205; Gaps 27;

QY 749 LCSLSEA--ICPFLGAL-----TPPAVTEFALTSGLAVIDLFLQMSAF 791
DB 602 LTSLSAFPFFFLAALVCLTMTWVEQRTQGTGKAFYSNG-SILKLYLVGSGA 660
QY 792 VALLSDSKROBASRLDCCVCKPOELPPPGGEGLLGF-----FKAYA 837
DB 661 SVI-----GVLGILIGFQFPNIIFNAYKSMYEMPSV 693
QY 838 PELLHWITRGVVLFLPLAFGVSL--YSWCHISVGLDQELALPKDSYLLDYFLNRYPE 895
DB 694 DIGFTWSYS--LLALFVALFCTTFTAYVACRAELRANAATMRPKAPKIGKIFLERISF 751
QY 896 VGAPVVF--VTTLGVNFSSEAGMAICSSAGC-----NNPSTQKIOY----- 936
DB 752 IWRNWFTSKVARNLFRYKQRMMLTVLGVAGCTALILTGFLGRNSISDIACQYGOIMK 811
QY 937 --ATBFPPEQSYLAIPASSWVDDFIDWLTSPSSCCRLYISGPNKDKFCPTVNSLNCNKM 994
DB 812 YDAAIYQDMS--APPAK--ETFDIMDNSI-----KSLAMSQTN-IEYVKSQ 857
QY 995 SITMGSRVPSVEQFHLYLPWFUNDPRNIKCPKGLAAYSTSVNLSDGOVLASRPMAYH- 1053
DB 858 SAQTTTSI-----IVPENLNEPNYIVLRD--RASHTTEKLTDGAIITEKLAKLFD 906
QY 1054 -KP-----LKNSQDYTEALRAARELANITADLRKVPGTDPAPEVPPYTIITVFEQYLT 1107
DB 907 VKPGDTITVKNANDKPOI-----KVSAITENYAHYIMTKAYYQQVFK 951
QY 1108 ILPE--GLFML-----SLCLVPTFA--VSCLLGLDLRSGLNLLS 1144
DB 952 EKPSYNLDLMLKDTSEKVESDFAEKLTDKAILNVTFNSNVSSL-----NETLDSLN 1005
QY 1145 IMVILVDTVGFVWALNDISNVLNLSVAVGMSVEFVSHITRSPAISTKPTWLBRAKEA 1204
DB 1006 IVIVVLITSAAALLAFVLYNLTN-INV-----SERIRE- 1037
QY 1205 TISMSAVPAGVAMTNLCILVLGLAKAQLIQIFPFRNLNLLITLGLLHGLV 1256
DB 1038 -----LSTIKVLGYPKE-VTMYYVRENIILTMGIAAGFI 1072

RESULT 79
US-60-752-355-4330
; Sequence 4330, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4330
; LENGTH: 463
; TYPE: PRF
; ORGANISM: Pseudomonas putida KT2440
US-60-752-355-4330

Query Match 1.5%; Score 103; DB 8; Length 463;
Best Local Similarity 20.0%; Pred. No. 11;
Matches 130; Conservative 66; Mismatches 183; Indels 270; Gaps 32;

QY 203 LNFQGDYTN---GLAPLDITF-----HLLPQAVGSGIQPLNEGVARCNESQDDVATCS 255
DB 1 MNTVSGDGNLAQKPRHVTMLSIAGIIGAGLFGVSG----- 37
QY 256 CODCAASCPAIPARQALDSTFVGLQMPGSLVLIILCSVFVAVVTTLLGVFRVAPARDKSK 315
DB 38 -----HAIAAAGPATIISYFVAGTL---VVLVNRMLGEMAV----- 70
```

```
QY 316 MVDPKKGTSLSKLSSTHTLIGQFFQCGWGTWVASWPLTILVLSVIPVVALAAGLVFTTEL 375
DB 71 -AHPDTG-----SFSIVA--DQAIRGWAGTYIGLWYFWVWVLPVPIEALAAGHVLNAM 120
QY 376 TTDPELWSAPNSQARSEKAFHQHGFPPFRFNQVILTPAPNRSSRYDSLLGLPKNFSGI 435
DB 121 FPQ-VDSW-----IFALASVLLA----- 138
QY 436 LDLDLLELLELLEQLRLHLQVMSPEAQRNISIQDICYAPLNPDPNTSLYDCCINSLLQYFQ 495
DB 139 -----GTLNF-----SVAKYGE 150
QY 496 NNRITLLLTANQTL-----MGQTSQVDWKDHFLYCANAPLTFKDGKTALALSCHADYG--AP 549
DB 151 PEFWPAILKVTAILGFIGLGPAAALMDW-----LPNREVSGLS-GLMAYGGPAP 198
QY 550 -----VPPFL-----AIGYKGDYSE--AEAL-----IMTP-----SL 576
DB 199 KQMSAVVGAFITVMPFSFIGTEAVTTAASESSDPSRNAKATRSVIRISTFYILSIFVII 258
QY 577 NNPAGDPRLAQAKLWBAFLSEMR---AFQRRMAGMFQVTFTAERSLEDEINRTAED 632
DB 259 SVVPWNPQLAVVGSYQRA--LEIMNIPNAAFMVDLVVLVAVTSCMNSSI----- 306
QY 633 LPIFATSYIVIFLYISLAL-----GSYSWSRVMVDSKATLGLGGVAVVVLGAVMAAMGPF 687
DB 307 -----YIASMFPESLAKRGDAPAFANLTKSVGPVRAAVFG---STLIGAAIAVLYNF 355
QY 688 SYLGRSLVILQVPPVLVSVGADNIFIP-VLEYQRLPRRPGEPREHVHIGRALGRVAPS 746
DB 356 APKG-----VPEFLASSGATALLVYVMAISQLRMR-----RRLERENTE 396
QY 747 M-----LLCSLSEATCFPL-GALTPM---PAVRT-FALTSGLAVIDLDEL 785
DB 397 LKRFMWLPPLYLTWAVIIFIAGALAVMMYTPHRAEVSSTLGLAIVISPL 445

RESULT 80
US-10-703-799B-246
; Sequence 246, Application US/10703799B
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: BGI-124CPCN
; CURRENT APPLICATION NUMBER: US/10/703,799B
; CURRENT FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: 09/603,208
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
```



```
QY 1154 GFMAWDISYNA-----VSLNLSVAGMSVEFVSHITRSPAIST----- 1193
Db 458 G-----PVDNNAQIAEMSGMDPAGAAIITSLDVGNTTK---AITKGIAIATAVLAAS 509
QY 1194 -----KPTWLERAKEATISMSAVFAGVAMTNLPGIILVGLAKAQLIQIFFEFLNLL 1245
Db 510 ALFGSFTDTVTTLRDAGALPEAGRGIVGGLNIA-YPDALV-GLTIGASV-VFLFS-GLA 565
QY 1246 IITLGLLHGLVFLPV 1260
Db 566 INAVGRAAGRVVLEV 580

RESULT 85
PCT-US06-07642-14490
; Sequence 14490, Application PC/TUS0607642
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462014340
; CURRENT APPLICATION NUMBER: PCT/US06/07642
; CURRENT FILING DATE: 2006-03-03
; NUMBER OF SEQ ID NOS: 26898
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14490
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from environmental sample
; NAME/KEY: DOMAIN
; LOCATION: (35)...(133)
; OTHER INFORMATION: MMPL family
; NAME/KEY: misc feature
; LOCATION: (1)..(304)
; OTHER INFORMATION: Xaa is any amino acid
PCT-US06-07642-14490

Query Match 1.5%; Score 102; DB 1; Length 304;
Best Local Similarity 24.0%; Pred. No. 7;
Matches 49; Conservative 39; Mismatches 78; Indels 38; Gaps 10;

QY 611 FOVTFTAERSLEDEINRTTAEDLPF---ATSVIVIELYISLALGSSYSSWRVWVDSKAT 667
Db 65 FOVLLAGTATINDTIKSKMSDMQKFVKLVLLIIAIFLFLPK----- 107
QY 668 LGLGGVAVVL-GAVMAAGFFSYLGIRSS--LVILQVVP--FLVLSVGAD--NIFIFVLE 720
Db 108 -RLSGVILPLFTVILAIITISLMAIFSPITVTVQILPSPFLAITIGASVHLLAIFYKD 166
QY 721 YQRLPRPGEPREVIHGRALGRVAPSMLLCSLSEAI---CFFLGALTPMPAVRTFALTSG 777
Db 167 YDF-----SKDKKASLKYALGHSLAIVMTSLTTAAGLWSFSFSELAPVADLGKFA-SSG 220
QY 778 LAVILDF-LLOMSAFVALLSLDSK 800
Db 221 ILLSLMFTLILLPALIATINIKPK 244

RESULT 86
US-60-658-984A-14490
; Sequence 14490, Application US/60658984A
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric
; APPLICANT: CHANG, Cathy
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564463014300/D2170-P1
; CURRENT APPLICATION NUMBER: US/60/658,984A
; CURRENT FILING DATE: 2005-03-04

PCT-US06-07642-14490

Query Match 1.5%; Score 102; DB 1; Length 304;
Best Local Similarity 24.0%; Pred. No. 7;
Matches 49; Conservative 39; Mismatches 78; Indels 38; Gaps 10;

QY 611 FOVTFTAERSLEDEINRTTAEDLPF---ATSVIVIELYISLALGSSYSSWRVWVDSKAT 667
Db 65 FOVLLAGTATINDTIKSKMSDMQKFVKLVLLIIAIFLFLPK----- 107
QY 668 LGLGGVAVVL-GAVMAAGFFSYLGIRSS--LVILQVVP--FLVLSVGAD--NIFIFVLE 720
Db 108 -RLSGVILPLFTVILAIITISLMAIFSPITVTVQILPSPFLAITIGASVHLLAIFYKD 166
QY 721 YQRLPRPGEPREVIHGRALGRVAPSMLLCSLSEAI---CFFLGALTPMPAVRTFALTSG 777
Db 167 YDF-----SKDKKASLKYALGHSLAIVMTSLTTAAGLWSFSFSELAPVADLGKFA-SSG 220
QY 778 LAVILDF-LLOMSAFVALLSLDSK 800
Db 221 ILLSLMFTLILLPALIATINIKPK 244

RESULT 87
US-60-752-355-22441
; Sequence 22441, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22441
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Woibachia endosymbiont of Drosophila ananassae
US-60-752-355-22441

Query Match 1.5%; Score 102; DB 8; Length 569;
Best Local Similarity 20.7%; Pred. No. 18;
Matches 111; Conservative 68; Mismatches 146; Indels 210; Gaps 27;

QY 276 FYLGQMPGSLVLIILCSVFAVVVITLLVGVFRVAPARDKSMVDPKGTSLSDKLSFSTHT 335
Db 9 FVYIKPNLSYFVIAFAVLFSAITLLFG-----RGLSNIID--SGT-----EHDFTTKL 56
QY 336 LLQGFQFGWGTWVASWPLILVLSVTPVVALAAGLVFTTELTTDPVLMAPSQAQSEKA 395
Db 57 L-----VAIIIVLAISLTA-----FTRLFYFI-----GSEKV 84
QY 396 FHDQHFQFPFRNTQVILTPAPNSVYRYSLLLPKPNFSGILDLLLELLELQLERHLQ 455
Db 85 I-----ARITYDL-----YSSITDL----- 99
QY 456 VMSPEAQRNISLQDICYAPLNPNTSLYDCCINSLLYQFQNNRTLLLTANQTLMGQTSQ 515
```

```
QY 1154 GFMAWDISYNA-----VSLNLSVAGMSVEFVSHITRSPAIST----- 1193
Db 458 G-----PVDNNAQIAEMSGMDPAGAAIITSLDVGNTTK---AITKGIAIATAVLAAS 509
QY 1194 -----KPTWLERAKEATISMSAVFAGVAMTNLPGIILVGLAKAQLIQIFFEFLNLL 1245
Db 510 ALFGSFTDTVTTLRDAGALPEAGRGIVGGLNIA-YPDALV-GLTIGASV-VFLFS-GLA 565
QY 1246 IITLGLLHGLVFLPV 1260
Db 566 INAVGRAAGRVVLEV 580

RESULT 85
PCT-US06-07642-14490
; Sequence 14490, Application PC/TUS0607642
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462014340
; CURRENT APPLICATION NUMBER: PCT/US06/07642
; CURRENT FILING DATE: 2006-03-03
; NUMBER OF SEQ ID NOS: 26898
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14490
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from environmental sample
; NAME/KEY: DOMAIN
; LOCATION: (35)...(133)
; OTHER INFORMATION: MMPL family
; NAME/KEY: misc feature
; LOCATION: (1)..(304)
; OTHER INFORMATION: Xaa is any amino acid
PCT-US06-07642-14490

Query Match 1.5%; Score 102; DB 1; Length 304;
Best Local Similarity 24.0%; Pred. No. 7;
Matches 49; Conservative 39; Mismatches 78; Indels 38; Gaps 10;

QY 611 FOVTFTAERSLEDEINRTTAEDLPF---ATSVIVIELYISLALGSSYSSWRVWVDSKAT 667
Db 65 FOVLLAGTATINDTIKSKMSDMQKFVKLVLLIIAIFLFLPK----- 107
QY 668 LGLGGVAVVL-GAVMAAGFFSYLGIRSS--LVILQVVP--FLVLSVGAD--NIFIFVLE 720
Db 108 -RLSGVILPLFTVILAIITISLMAIFSPITVTVQILPSPFLAITIGASVHLLAIFYKD 166
QY 721 YQRLPRPGEPREVIHGRALGRVAPSMLLCSLSEAI---CFFLGALTPMPAVRTFALTSG 777
Db 167 YDF-----SKDKKASLKYALGHSLAIVMTSLTTAAGLWSFSFSELAPVADLGKFA-SSG 220
QY 778 LAVILDF-LLOMSAFVALLSLDSK 800
Db 221 ILLSLMFTLILLPALIATINIKPK 244

RESULT 86
US-60-658-984A-14490
; Sequence 14490, Application US/60658984A
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric
; APPLICANT: CHANG, Cathy
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564463014300/D2170-P1
; CURRENT APPLICATION NUMBER: US/60/658,984A
; CURRENT FILING DATE: 2005-03-04

PCT-US06-07642-14490

Query Match 1.5%; Score 102; DB 1; Length 304;
Best Local Similarity 24.0%; Pred. No. 7;
Matches 49; Conservative 39; Mismatches 78; Indels 38; Gaps 10;

QY 611 FOVTFTAERSLEDEINRTTAEDLPF---ATSVIVIELYISLALGSSYSSWRVWVDSKAT 667
Db 65 FOVLLAGTATINDTIKSKMSDMQKFVKLVLLIIAIFLFLPK----- 107
QY 668 LGLGGVAVVL-GAVMAAGFFSYLGIRSS--LVILQVVP--FLVLSVGAD--NIFIFVLE 720
Db 108 -RLSGVILPLFTVILAIITISLMAIFSPITVTVQILPSPFLAITIGASVHLLAIFYKD 166
QY 721 YQRLPRPGEPREVIHGRALGRVAPSMLLCSLSEAI---CFFLGALTPMPAVRTFALTSG 777
Db 167 YDF-----SKDKKASLKYALGHSLAIVMTSLTTAAGLWSFSFSELAPVADLGKFA-SSG 220
QY 778 LAVILDF-LLOMSAFVALLSLDSK 800
Db 221 ILLSLMFTLILLPALIATINIKPK 244
```



```
Db 100 --OPSFENTGVQDVISALIT--DTSVLQSIINSSL-----LTIILRNFFVLIGSVAM 147
Qy 516 VDKOHFLYCANAPITFKDGTALALSCMADYGAPEVPPFLAIGYKGYKDYSEBALIMTFS 575
Db 148 L-----LY-TNQLT-----AYAAAIIPIL-----LIIVTS 172
Qy 576 LNNYPAGDPRLAQAKWEBA-FLEENRAPORRMAGMPQVTFPTAERSLEDEINRTT-AEDL 633
Db 173 LGKVRSHARFAQDKLSELASFSE-----NFRSIVITIKSPVLEENKTRFKEYL 222
Qy 634 PIPATSYIIVFLYISLALSGYSWSRWMDVDSKATLGLGGVAVVLGAVMAAMGFFSYLGR 693
Db 223 NSVSKSYVKLVLL-----RAIL-----VTLVITCVIGSLVLLFFFGIK 260
Qy 694 ----SSLVILQVPPFLV----LSVGA-----DNIFIP-----VLEYORLPRRGEPR 732
Db 261 EVLSNNITIGELSSFFVYVYALAGAINNLSIDLQGLGIVERLFEFNKMKSSIVDPD 320
Qy 733 EVHIGRALGRVAPSMMLCSLSBAI-----CFPLGALTMPMPAVR--TFALTSGLA 780
Db 321 D-----PIKICSVQKGISPGVTFPFVESQSDKPKALNNVVSFSIEAGQAV 363
```

RESULT 88

```
US-10-498-451-1750
; Sequence 1750, Application US/10498451
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-ru
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Ghosh, Malabika
; APPLICANT: Xue, Aidong J.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Weng, Gezhi
; APPLICANT: Xu, Chongjun
; APPLICANT: Boyle, Bryan J
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/498,451
; CURRENT FILING DATE: 2004-06-09
; PRIOR APPLICATION NUMBER: US 60/339,739
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 10/128,558
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/339,453
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/365,384
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/365,091
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/372,615
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 60/376,045
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 60/372,381
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 3134
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1750
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-498-451-1750

Query Match
Best Local Similarity 19.7%; Pred. No. 1.2e+02;
Matches 174; Conservative 108; Mismatches 257; Indels 346; Gaps 49;

Qy 501 LLLLTANQTLMGQTSQVDMKOHFLYCANAPLT--FKDGTALALSCMADYGAPEVPPFLAIGG 558
Db 875 LLPQNTCLMYLLEQIDMHGFGCGTAATSPLTAVFSLSRSLLAAALL-YG-----FCLGA 927
Qy 559 YKKG-----DYSEAEALIMTFSLN-NYPAGDP-----RLAQAKLM-----EFAFLEMRA 602
Db 928 IKTPMPHQHPVFLFSVFCGLLVALSYHLSRQSDPTVLSLRSKLFPPELEERSLETARA 987
Qy 603 FORMAGMPQVTFPTAERSLEDEINRTTARDLPFATSYIIVFLYISLALSGYSWSRWVMV 662
Db 988 -----EPDPLPDKMRQSVREVL-----HSDLMVCVVIATVLAISA-STVFI 1029
Qy 663 DSKATLGLGGVAVVLGAVMAAMGFFSYLGRISLSVLQVPPFLVLSV----- 709
Db 1030 ALKSVLGF-----VLVALAGAVGFTHYLLPQ---LRKQLPWFCLSQPVLPKLEYSQYEV 1081
Qy 710 -GADNIFIVLEYQRLPRRGEPRVHIGRALGRVAPSMMLCSLS-----EAIQF 758
Db 1082 RGAQVMMFPEKLYAGL-----QCVEKYL-----IYPVVLNALTVDAAHTVSHDPKCYF 1130
Qy 759 FLGALTMPMPAVRTFALTSGLAVIDFLQMSAPFALLSDSKRQASRLDVCCKVPQEL 818
Db 1131 YCRAL-----LMTVAGLKLRSAF-----CC----- 1151
Qy 819 PPPQGGEGLLGFFQKAVAPFLLHWTIRGWILLFLALFGVSLYSCHISVGLDQELALP 878
Db 1152 -PPQQ-----YLTLAFTVLLF-----HFDYPRLSQG----- 1176
Qy 879 KDSYLLDFLNRYPFVGCAPVYFTTILGYNFSSBAGMNAICSSACNNPSTFKIQYAT 938
Db 1177 ---FLLDYFL-----MSLLCSKL-----WDLLYKLRFVL 1202
Qy 939 EPPEQSYLAIPASSWVDDFIDMLTPSSCCRLYISGPNKDKPCPSTVNSLCLNCKMSITM 998
Db 1203 -----TYIAPWQITWGSAPFAQAQPPACATY-----PACV----- 1232
Qy 999 GSVRPSVEQPHKYLEPWFNLDRENIKCPKGGLAAYSTSVNLTSQGVLASRFRMAYHKPLKN 1058
Db 1233 -----WAL-----LSGL--PSTPLNPLIGSAVF---IMSYARPLKF 1263
Qy 1059 -SDYTEALR-----AARELAA-----NITADLRKVPCTDPAFEVFPY 1095
Db 1264 WERYKESMSRVHPLFTFLANRLVALSLNSTRKVDHNSRTRVLTQDRNFGAD-----DN 1317
Qy 1096 TITNVFYEQYLTILPEGLFMLSIC--LV-----PTFAVSCLLGLDLRSLGLMLLSI 1145
Db 1318 NLNSIFYEHLTRSLQH-----TLCGDVLGRWGNVGPQ---DCFVLASDYLNALVHLIEV 1369
Qy 1146 VMILVDYVGFMAIMDISYNAVSLINL-VSAVGMSEY-----FVSHITR--SFAISTK 1194
Db 1370 GNGLVT----FQLRGLEFRGTVCQQQREVEAITEGVEDEBGGCCCPGHLPRVLSFNAAG 1425
Qy 1195 PTWLERAKBATISMGSAVPAVAMTNLPGLIVGLAKAQLIQIFPFLMLLTILGLLHG 1254
Db 1426 QRWL--AWEVITAS--KYVLEGYSISDNNA-----ASMLQVFDLR-KILIT----- 1465
Qy 1255 LVFLPVILSYGPDVNPALALEQKAEBAVAAMVASCNHPSPRV 1299
Db 1466 -YYVKSIIYY-----VSRPKLEVWLSHEGITAAL-----RPRV 1499
```

RESULT 89	
US-60-781-306-41	
; Sequence 41, Application US/60781306	
; GENERAL INFORMATION:	
; APPLICANT: MESRI, Mehdi et al.	
; TITLE OF INVENTION: PROSTATE DISEASE TARGETS AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: CLO01683	
; CURRENT APPLICATION NUMBER: US/60/781.306	
; CURRENT FILING DATE: 2006-03-13	
; NUMBER OF SEQ ID NOS: 547	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 41	
; LENGTH: 634	
; TYPE: EXT	
; ORGANISM: Homo sapiens	
US-60-781-306-41	
Query Match 1.5%; Score 101.5; DB 8; Length 634;	
Best Local Similarity 19.3%; Pred. No.23;	
Matches 105; Conservative 74; Mismatches 161; Indels 205; Gaps 28;	
Qy	522 FLYCANAPLTFKOGTALALSCMADYGA PV-----FPFLAIGGYKGYKDYSEAEALIMT 573
Db	72 FLIAALA-----SVNAGLCYAEFGARVPKTSAYLYTYTVVG-----ELWAFITG 116
Qy	574 FSL-NNYPCAGDPLRAQAKLWEAEFLKEWRAFORRMAGMFQVTFPAERSLDEINRTTAED 632
Db	117 WNLTLSYVIGTSSVARA--WSGTDELL-----SKQIGQLRTYF-----RMNYTGLAE 163
Qy	633 LP-----IFATSYVIVFLVISLA---LGSYSW-----657
Db	164 YPDFFAVCLILLAGLLSFGKESAWNKVFTAVNILLVLFVMVAGFVKGNAVWKISEE 223
Qy	658 --SRVWDSK-----ATLGLGG-----VAVVLGAVMAAMGFFSY-----689
Db	224 FLKNISASAREPPSENGTSTYGAGGFMYPGFTGTLAGAATCFYAFVGFDCIATTGEBVN 283
Qy	690 -----LGIRSSSILVQVPFLVLSVCADNIFIFVLVEYQRLPRPGEP---REVHIGRAL 740
Db	284 PQKAIPIGIVTSLVCFMAYF-----GVSAALILMPEYLLDEKSPLPVAFEVYVGWGP 338
Qy	741 GRVAPSMLLCSLSEAICFFLGAITPMAV-----RT---FALTS 776
Db	339 YVVAAGS-LCALSTSL---LGSIFPMPRVIYAMAE DGLLFKCLAIQINSKTKTPIATLSS 394
Qy	777 G-LAVILDFLLQWSAFVALLSLDS-----KQOEASRLD 808
Db	395 GAVAALMAFLFDKALVDMMSIGTLMAYSVAACVILRVTSKESQVTLMLQRGFSMRT 454
Qy	809 VCCCVKPEQLPPPGQGE--GLLLGFFQ-----KAVAPFLHLWITR-----GVVLLFLA 855
Db	455 LFC---PSLLPTQOSASLSVFLVGLFLAFLVLGLSVLTYGVGHAI TRLEAWSLALLALFLV 511
Qy	856 LFGVSLYSMCHISVGLDQELA-----LPKDSYLLDYPFLF-----NRVFEVG 897
Db	512 LP-VAIVLTIWRQPNQKQVAFMVPFLPFLPAFSILVNIYIMVQLSADTWVRFESINWAIG 570
Qy	898 APVYF 902
Db	571 FLIYF 575
RESULT 90	
US-10-917-503B-12928	
; Sequence 12928, Application US/10917503B	
; GENERAL INFORMATION:	
; APPLICANT: OTA, TOSHIO	
; APPLICANT: ISOGAI, TAKAO	
; APPLICANT: NISHIKAWA, TETSUO	
; APPLICANT: HAYASHI, KOJI	
; APPLICANT: SAITO, KAORU	
; APPLICANT: YAMAMOTO, JUNICHI	

```
; NUMBER OF SEQ ID NOS: 2517  
; SEQ ID NO 2249  
; LENGTH: 397  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-533-519-2249
```

		Query Match	1.5%; Score 101; DB 6; Length 397;
		Best Local Similarity	26.9%; Pred. No.13;
		Matches	57; Conservative 35; Mismatches 74; Indels 46; Gaps 12;
Qy	1101 FYEQLTI-----LPBGLPMLSLCLVPTF-AVSCLLGLDLRLSGLINLISIVMIL----	VD	1151
Dd	24 YFESYLAVASVP-----MLCLVANELLNVRVAHVIRVLASLTWILAIFMWITALVKVD	78	
Qy	1152 TVGPMALWDISTSYNAVLSINLVSAVGMSVEPVSHI-----TRGFATSTKPTWLERAKEATIS-	1207	
Dd	79 T-----SSWRGRFFAFTVCVMVILSGASTVFSSSIYGMTGSPPM-----RNSQUALIS	126	
Qy	1208 --MGSAVPAGVAMTNLPAIL-VLGAKAQLIQIPFPFLRNLLITILGLHLGVLFPLVTLSY	1264	
Dd	127 GAMGCTVSVASLVLDLAASDVNRNALA-----FFLTATITELV---LCMGLYLLLSRLEY	178	
Qy	1265 VGPDVNPALAL-----EQKRABEAFAAAVMVAS	1291	
Dd	179 ARYYMRPVLAHHVFSGEHEELPODSLSPAPSVAS	210	

```

RESULT 92
US-10-206-921A-260
; Sequence 260, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: 39780-3430R1C515
; CURRENT APPLICATION NUMBER: US/10/206,921A
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 260
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-206-921A-260

```

[illegible]

```

Qy 1208 --MGSVAFGAVGAVMTNPGIIL-VGLAKAQLIQIPFFPLNLIITLGLLHGLVFLVPLVILSY 1264
Db 205 GANGGTGSVASVSLVDLSAASDVNSALA-----FFLTATIFLV---LCMGLYLLLSRLEY 256
Qy 1265 VGPDVNPALAL-----EKKRAEEAAVAVMVAS 1291
Db 257 ARYMREPLVAHVPSGEEELPQDSLSAPSASV 288

RESULT 93
US-10-184-614A-260
; Sequence 260, Application US/10184614A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/184,614A
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 260
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-614A-260

```

	Query Match	1.5%;	Score 101;	DB 6;	Length 475;
	Best Local Similarity	26.9%;	Pred. No. 17;		
	Matches	57;	Conservative	35;	Mismatches 74; Indels 46; Gaps 12;
Qy	1101 FYEQLTI---	LPBGLPMLSLCVPFF-AVSCLLGLGLDLSGILNLLSI	VMIL----	VD	1151
Db	102 YFBSYLAVASTVPS----	MLCVAMFLNNRVAVHIVLASITVILAI	FMVITALKVD	156	
Qy	1152 TVGFMALWDISYNAVSLINILSVAGMSVEFVSHI---	TRSFATSTKPTWMLERAKEATIS-		1207	
Db	157 T-----	SSWTGTFPAVITVCVILSGASTVPSSIYGWTSFPM-----	RNSQALISG	204	
Qy	1208 --	MGSVAFGAVGMVNNPGIL-VGLAKAQLIQI	FPFRLNLLITLGLLHGIVFLPVILSY	1264	
Db	205 GAGMGGTVSASVLDLAASSDVRNSALA-----	FFLTATIFLV---LCMGYLLLSRLSY	256		

Qy 1265 VGPDPNPALAL-----EQKRAEEAAVAVMVAS 1291
 Db 257 ARYYMRPVLAAHVFSGEEELPQDLSAPSVAS 288

RESULT 94

US-10-184-615A-260
 ; Sequence 260, Application US/10184615A
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pap, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C1
 ; CURRENT APPLICATION NUMBER: US/10/184,615A
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059266
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063120
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063121
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063486
 ; PRIOR FILING DATE: 1997-10-21
 ; PRIOR APPLICATION NUMBER: 60/063540
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063541
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063544
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063564
 ; PRIOR FILING DATE: 1997-10-28
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 260
 ; LENGTH: 475
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien

US-10-184-615A-260

Query Match 1.5%; Score 101; DB 6; Length 475;
 Best Local Similarity 26.9%; Pred. No. 17;
 Matches 57; Conservative 35; Mismatches 74; Indels 46; Gaps 12;
 Qy 1101 FFEQYLTIT---LPEGLFMLSCLVPTFP-AVSCLLGLDLRSGLNLLSIWMIL-----VD 1151
 Db 102 YFESYLAVASTVPS-----MLCLVANFLVNVRVAHVIRVLASLTIVLAIFMVTALVKVD 156
 Qy 1152 TVGFMALWDISYNAVSLINLVSAGVMSVFVSHI---TRSFATSTKPTWLERAKEATIS- 1207
 Db 157 T-----SSWTRGFPAVTIVCMVILSGASTVFSSSYGMTGSFPM-----RNSQALISG 204
 Qy 1208 --MGSVAFAGVAMTNLPGL-VLGLAKAQLIQIFFFRNLNLLITLLGLHGLVFLPVILSY 1264
 Db 205 GAMGGTTSVAVASLVDLAASSDVNSALA-----FFLTATIFIV---LCMGLYLLLSRLEY 256
 Qy 1265 VGPDPNPALAL-----EQKRAEEAAVAVMVAS 1291
 Db 257 ARYYMRPVLAAHVFSGEEELPQDLSAPSVAS 288

RESULT 95

US-60-732-162-1676
 ; Sequence 1676, Application US/60732162
 ; GENERAL INFORMATION:
 ; APPLICANT: Belouchi, Abdelmajid
 ; APPLICANT: Raelson, John V
 ; APPLICANT: Bradley, Walter E
 ; APPLICANT: Paquin, Bruno
 ; APPLICANT: Fournier, Helene
 ; APPLICANT: Nguyen-Huu, Quynh
 ; APPLICANT: Croteau, Pascal
 ; APPLICANT: Allard, Rene
 ; APPLICANT: Debrus, Sophie
 ; APPLICANT: Eerdewegh, Paul V
 ; APPLICANT: Little, Randall D
 ; APPLICANT: Keith, Tim
 ; APPLICANT: Segal, Jonathan
 ; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
 ; FILE REFERENCE: 059908-5010-PR
 ; CURRENT APPLICATION NUMBER: US/60/732,162
 ; CURRENT FILING DATE: 2005-11-02
 ; NUMBER OF SEQ ID NOS: 4417
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 1676
 ; LENGTH: 475
 ; TYPE: PRT
 ; ORGANISM: Homosapiens
 ; US-60-732-162-1676

Query Match 1.5%; Score 101; DB 8; Length 475;

Best Local Similarity 26.9%; Pred. No. 17;
 Matches 57; Conservative 35; Mismatches 74; Indels 46; Gaps 12;
 Qy 1101 FFEQYLTIT---LPEGLFMLSCLVPTFP-AVSCLLGLDLRSGLNLLSIWMIL-----VD 1151
 Db 102 YFESYLAVASTVPS-----MLCLVANFLVNVRVAHVIRVLASLTIVLAIFMVTALVKVD 156
 Qy 1152 TVGFMALWDISYNAVSLINLVSAGVMSVFVSHI---TRSFATSTKPTWLERAKEATIS- 1207
 Db 157 T-----SSWTRGFPAVTIVCMVILSGASTVFSSSYGMTGSFPM-----RNSQALISG 204
 Qy 1208 --MGSVAFAGVAMTNLPGL-VLGLAKAQLIQIFFFRNLNLLITLLGLHGLVFLPVILSY 1264
 Db 205 GAMGGTTSVAVASLVDLAASSDVNSALA-----FFLTATIFIV---LCMGLYLLLSRLEY 256
 Qy 1265 VGPDPNPALAL-----EQKRAEEAAVAVMVAS 1291
 Db 257 ARYYMRPVLAAHVFSGEEELPQDLSAPSVAS 288

RESULT 96

US-60-752-355-7829
 ; Sequence 7829, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60/752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 7829
 ; LENGTH: 556
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; US-60-752-355-7829

Query Match 1.5%; Score 101; DB 8; Length 556;

Best Local Similarity 16.9%; Pred. No. 21;
 Matches 107; Conservative 92; Mismatches 206; Indels 228; Gaps 27;
 Qy 499 TLLLLLTANTQTLMGQT-----SQVDWKDHFYCANAPLTFKDGDTALALSCNADYGAPVFPPL 554

```
Db 71 TWMLASTIVPANGTPEGDKARVQISFLMSGINTLLQTLVGRTRLPVTVMNASP-AFVVPVL 129
QY 555 AIGGYKGYDYSAEALIMTFSLNYPAGDPRL-----AQAKLWBEAFLEEM-----R 601
Db 130 SI-----AKDFEQ-----NNYASSHQRTHTWRAIQGALIVASILNMLIGSTWIG 175
QY 602 AFQRMA-----GMQVTPTAERLEBINTTARDLPFATSVIVIFLYIS 648
Db 176 AFARFSPVIMTPVVCVGLGLFALGF-----PQVGKCVIGLPLMLILA-VVQQVYP 227
QY 649 LALGSYSSSRVWVDSKATLGLGGVAVLGVMAAGPFESYLGRS-----SLVILQV 701
Db 228 YFHHYHERITLFEKYSLLLCIGIVMAFAALITAGAYNHVSLTKTQQCHRTDKSYLISS 287
QY 702 VPFLVL-----SVGADNIFIFVLEYQ-----RLPRRPGEPREV 734
Db 288 APWIKIPYPFQNGTPIFTAGHSFGMGA---VLVSAPFSTGNHFAFATLAGATPPASV 343
QY 735 HIGRALGRVAPSMML-----CSLSEAI CFFLGALTPMPAVRTFALTSGLAVILDPL 785
Db 344 -LSRSVGLQIGIFLEGIFGAPAGSSVSVENIGLLG-LTKVGSRRVIQISTGFMIFFSIF 401
QY 786 LQMSAFVALLSLDSKQEARLDVCCVKPQELPPPGQEGLLGFFQKAYAPFLHMIT 845
Db 402 GKGFAPFA-----SIPLP-----414
QY 846 RGWLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYPEVGAPVYVVT 905
Db 415 --IPAAIFILGIV-----NAVGSYMQFNKSMRNIYIIGSLFGISVPQYF--- 463
QY 906 LGYNFSSEAGMAICSSAGCNFFSFTQKIQYATEPPEQSYLEIPASSWVDDFIDMLTSS 965
Db 464 --HEVTASASTGPARTNAG-----WFDNIIN-----487
QY 966 CRLYISGNKDKFCPESTVNSLNCNKMISITMGSVRPSVEQFHKY-----LPW---FLN 1017
Db 488 --TVFASGPT-----VSLIVASILDNTLEFRGYENDRGLEPWFMPFLH 527
QY 1018 DRPNIKPKGGLAAYSTSVNLTSDGOVLASREM 1050
Db 528 RRGKYSDFRND-EFYSFPIRV---HDVIPSREFL 556
```

RESULT 97

```
US-11-174-307B-3734
; Sequence 3734, Application US/11174307B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; TITLE OF INVENTION: USEFUL FOR MODIFYING PLANT CHARACTERISTICS
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 3734
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: PTR2; Pfam Description: POT family
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
```

RESULT 98

```
PCT-US05-46487-19
; Sequence 19, Application PC/TUS0546487
; GENERAL INFORMATION:
; APPLICANT: Sheets, Joel J.
; APPLICANT: Ni, Weiting W.
; APPLICANT: Larrinua, Ignacio M.
```

```
; OTHER INFORMATION: GI Number: 16226701; NR Description: AT3g47960/T17F15_170
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 6730698; NR Description: Similar to peptide
; OTHER INFORMATION: transport proteins [Arabidopsis thaliana]
; OTHER INFORMATION: >gi21436191|gb|AA51383.1| putative peptide transporter protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11933397; NR Description: nitrate transporter
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11933400; NR Description: nitrate transporter
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11933414; NR Description: nitrate transporter
; OTHER INFORMATION: GI Number: 11933414; NR Description: nitrate transporter
US-11-174-307B-3734

Query Match 1.5%; Score 101; DB 7; Length 607;
Best Local Similarity 19.6%; Pred. No. 24;
Matches 116; Conservative 72; Mismatches 219; Indels 184; Gaps 22;

QY 130 SPNQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAATLA--VGT 187
Db 86 SVNAATLLNVFS---GTSNLTATVFG-----AVSDTY--LGRVTTTAAATMSSFIGM 132
QY 188 MCGVYGSAL-----CNAQRWLNFGQDTGNGIAPLDITPHLLBEGQAVGSGIQLNREGV 240
Db 133 LITLTAATHLTHLPACNASKQHCEGPTGSLAAAILVSFPFL---VIGAG-----GI 182
QY 241 ARCNSQDDVATCSQDCQCAAPAIARPOALDS-----TPVLQMPGSLVL 287
Db 183 RPCNLAFGAQD-----FNPHTADGRGRIASFNNYFTFTVAMMLSATVI 227
QY 288 IILICSV----FAVVTILLVGRVAPARDKSKWDPKKGTSLSLDSKLSFSTHTLLGQFQG 343
Db 228 IYLSQNVNVALGLAVPALMGLSCA-----VFFWG 257
QY 344 WGTWVASWPLTILVLSVIPVVALAAGLVFTLTTPDVELWSAPN-SQARSEKAFHDQHFQ 402
Db 258 TRLYVRVRPEGSPFTSFAQLVVAARKHRIRARGDAELFDPHQSKLVSKLAYTDQ--- 314
QY 403 PFRTNQVILITAPNRSYRYSLLILGPKNFSGILDLILLELELQERLRLHLOVMSPEAQ 462
Db 315 -FACLDKAAVRTPD-----DALCIDGKTPADPWRLCTVQVEVKCLARIIPVWSSGIV 367
QY 463 RNISLQDI-CYAPLNPDTSLYDCCINSLLQVFQNNRTLLLTANQTLMGQ----- 512
Db 368 FFIIVLTQGTIV-----VLQAAQMDRIRISKSSSFQIQGSPVFMQLAL 411
QY 513 TSQVDMKDHFLYCANAPLTFKD-----GTALALSCMADYGAAPVFPFLAIGGYKGD 563
Db 412 TMIIPVYDRFVVVPALRRFTKREGITLQRIQVGLASV-----ATMVVSAAVEQRRKI 466
QY 564 YSEABA--LIMTFSLNYPAGDPRLAQAQLWEEAFLEENRAPQRMAQHWQVFTTAERSL 621
Db 467 GSSMSCFWLVPQQLLAGLSEAFGAIGQIEFYRQFPENRS-----VAG----- 510
QY 622 EDEINRTTAEDLPFATSVIVIFLVSIALGSYSSSRVWVDSKATLIGGG 672
Db 511 -----ALYFLGFAMASYASGLMVMVVRATRGRRG 540
```

```
; APPLICANT: Bevan, Scott A.
; TITLE OF INVENTION: Second Toxin Complex from Xenorhabdus nematophilus
; FILE REFERENCE: 63566PROV
; CURRENT APPLICATION NUMBER: PCT/US05/46487
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1016
; TYPE: PRF
; ORGANISM: Xenorhabdus nematophilus
PCT-US05-46487-19

Query Match      1.5%; Score 101; DB 1; Length 1016;
Best Local Similarity 18.3%; Pred. No. 53;
Matches 162; Conservative 83; Mismatches 229; Indels 410; Gaps 42;

Qy 41 KNPELGSGLMTLSNV-----SCLS-----NTPARKITGDHLILLQKICPRLY--TGNP 86
Db 124 ESDNLPGRLLTITEQVKGACITERLIWSGNTPAEK--GNN---LAGQCVVHYDPTGMN 178
Qy 87 TQACCSAKQLVSLASISITKALLTRCPACSDNFVNLHCHNTCSPNQSLFINVTRVAQLG 146
Db 179 QTNSIS---LTSI---PLSITQQLK-----DDSEADWHGMD-----209
Qy 147 AGQLPAVVAYEAFYQHSFAEQSYDCSRVRVPAATAATLAVGTMCVGYGSALCNAQR-----201
Db 210 -----ESGWNALAPESFTSVS-----TTDATGTVL--TSTDAAGNKQRIADV 251
Qy 202 -----WLNFGQ-----DTGNGLAPLDITFHLLEP-----225
Db 252 AGLQGSWLAKGQEQVIVKSLTVSAASQKLREHGNGI---VTTYTEPETQRVIGI 307
Qy 226 -----GOAVGSGI-----QPLNEGVARCNESEQ-----247
Db 308 KTERPSGHAAGEKILQNLRYEYDVGVLKSTNDAEITFRWRNQKIVPENTYTYDSLYQL 367
Qy 248 ---GDDVATCSODCAASCAPAIARPOALDSTFYLG-----QMPGSLVLIILCSVFA 296
Db 368 VSVTGREMANIGRQNKQIPALP-----IDNNTYNTYRTYDRGGLNTRI-----414
Qy 297 VVTILLVGRVAPARDKSNQVDPKKGTSLSKLSFSTHTLLGQFFQGGTGWASWPLTIL 356
Db 415 -----RHNS-----PITGNNYTNMTVSDH-----434
Qy 357 VLSVIPVVALAAGLVFTLTTPD--VELWSAPNSQARSEKAFDHQFGFPFRTNQVIL--412
Db 435 -----SNRAVLEELAQDPTQVDMLFTPGGHQTRLVPGQDLFWTPRDELOQVILVN 484
Qy 413 ---TAPNRSSVRYDS-----LLLGPK-----NPSGILDLDLLE 443
Db 485 RENTTPDQEFYRYDADSQRVKTHTKGTGNSQIORTLYLPELEWRTTYSG-----NTLKE 540
Qy 444 LLEL-----QERLRLHQ-----455
Db 541 FLOVITVSGSGAQVRVLHWETGKPADISNDQLRYSYGNLIGSSGLELSDGQIISQBEY 600
Qy 456 -----VMSPEAQRNISLDQICYAPLNPDNTSLYDCCINSLLQYFQNNRLLLLTANQT 508
Db 601 YPGGTAVWAARSQSEADYKTVRYSGKERDATGL-----YYYGYRYQSWTGRWL 650
Qy 509 LMGQTSQVWDKDFHLYCANAPLTFKDGTAALASCADYCAPVPFPAIGGYGKD--YSEA 567
Db 651 SYDPAGEVDGLNLFMRNNPVIFFSDSGR-----FPGQGLAWIGKKAYRKA 698
Qy 568 EALIMTFSLNNYPAGDPRLAQAKLWEEAFLEBMRAPFORMMAGMFOVTTAERSLSDEINR 627
Db 699 -----VNIITEHLLLEGASFD-----TEL-----KLNK 721
Qy 628 TTAEDLPFIPATSYIVIFLYISLA--LGSVSSWSRVMDVSKATIGLGGVAV---VLGAVMAA 683
Db 722 ---GLRTFVLGVGVASLVGKAAITAGASPW-----GIVMAIGGFVSGAV---763
```

```
Qy 684 MGPF-----SYLGIRSLVILQVVPFLVLSVGADNIP 715
Db 764 MGPFANNISEKIGEVLSYL--TRKRSVPVQVGAFTVTSLSALP 806

RESULT 99
US-11-314-892-19
; Sequence 19, Application US/11314892
; GENERAL INFORMATION:
; APPLICANT: Sheets, Joel J.
; APPLICANT: Ni, Weiting W.
; APPLICANT: Larrinua, Ignacio M.
; APPLICANT: Bevan, Scott A.
; TITLE OF INVENTION: Second Toxin Complex from Xenorhabdus nematophilus
; FILE REFERENCE: 63566PROV
; CURRENT APPLICATION NUMBER: US/11/314,892
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1016
; TYPE: PRF
; ORGANISM: Xenorhabdus nematophilus
US-11-314-892-19

Query Match      1.5%; Score 101; DB 6; Length 1016;
Best Local Similarity 18.3%; Pred. No. 53;
Matches 162; Conservative 83; Mismatches 229; Indels 410; Gaps 42;

Qy 41 KNPELGSGLMTLSNV-----SCLS-----NTPARKITGDHLILLQKICPRLY--TGNP 86
Db 124 ESDNLPGRLLTITEQVKGACITERLIWSGNTPAEK--GNN---LAGQCVVHYDPTGMN 178
Qy 87 TQACCSAKQLVSLASISITKALLTRCPACSDNFVNLHCHNTCSPNQSLFINVTRVAQLG 146
Db 179 QTNSIS---LTSI---PLSITQQLK-----DDSEADWHGMD-----209
Qy 147 AGQLPAVVAYEAFYQHSFAEQSYDCSRVRVPAATAATLAVGTMCVGYGSALCNAQR-----201
Db 210 -----ESGWNALAPESFTSVS-----TTDATGTVL--TSTDAAGNKQRIADV 251
Qy 202 -----WLNFGQ-----DTGNGLAPLDITFHLLEP-----225
Db 252 AGLQGSWLAKGQEQVIVKSLTVSAASQKLREHGNGI---VTTYTEPETQRVIGI 307
Qy 226 -----GOAVGSGI-----QPLNEGVARCNESEQ-----247
Db 308 KTERPSGHAAGEKILQNLRYEYDVGVLKSTNDAEITFRWRNQKIVPENTYTYDSLYQL 367
Qy 248 ---GDDVATCSODCAASCAPAIARPOALDSTFYLG-----QMPGSLVLIILCSVFA 296
Db 368 VSVTGREMANIGRQNKQIPALP-----IDNNTYNTYRTYDRGGLNTRI-----414
Qy 297 VVTILLVGRVAPARDKSNQVDPKKGTSLSKLSFSTHTLLGQFFQGGTGWASWPLTIL 356
Db 415 -----RHNS-----PITGNNYTNMTVSDH-----434
Qy 357 VLSVIPVVALAAGLVFTLTTPD--VELWSAPNSQARSEKAFDHQFGFPFRTNQVIL--412
Db 435 -----SNRAVLEELAQDPTQVDMLFTPGGHQTRLVPGQDLFWTPRDELOQVILVN 484
Qy 413 ---TAPNRSSVRYDS-----LLLGPK-----NPSGILDLDLLE 443
Db 485 RENTTPDQEFYRYDADSQRVKTHTKGTGNSQIORTLYLPELEWRTTYSG-----NTLKE 540
Qy 444 LLEL-----QERLRLHQ-----455
Db 541 FLOVITVSGSGAQVRVLHWETGKPADISNDQLRYSYGNLIGSSGLELSDGQIISQBEY 600
Qy 456 -----VMSPEAQRNISLDQICYAPLNPDNTSLYDCCINSLLQYFQNNRLLLLTANQT 508
Db 601 YPGGTAVWAARSQSEADYKTVRYSGKERDATGL-----YYYGYRYQSWTGRWL 650
```

QY 509 LMGQTSQVDWKDHFLYCANAPLTFKDGDTALALSCMADYGAPVFPFLAIGYKGD-YSEA 567
Db 651 SVDPAGEVDGLNLFMRNRPVFSDDGR-----PPGQGVLAWIGKKAYRKA 698
QY 568 EALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVFTFAERSLEDEINR 627
Db 699 -----SNRAVLLELAQDPTQVDMLFTEGQHQLTRVPCQDLFWTPRDELOQQVILVN 484
QY 628 TTAEDLPIPATSYIVIFLYISLA-LGSYSSWSRVMVDSKATLGLGVAV---VLGAVMAA 683
Db 722 ----GLRTFVLGVGVASLGVAATAGASPM-----GIVGAALGGFVSGAV--- 763
QY 684 MGFF-----SYLGRSSSLVILQVVPFLVLSVGADNIF 715
Db 764 MGFFANNISEKIGEVLSYL-TRKRSVPVQVGAFVVTSLVTSALF 806

RESULT 100

US-11-375-551-16
; Sequence 16, Application US/11375551
; GENERAL INFORMATION:
; APPLICANT: Bintrim, Scott
; APPLICANT: Mitchell, Jon
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Apel-Birkhold, Patricia
; APPLICANT: Schafer, Barry
; APPLICANT: Bevan, Scott
; APPLICANT: Young, Scott
; APPLICANT: Guo, Lining
; TITLE OF INVENTION: Xenorhabdus TC Proteins and Genes for Pest Control
; FILE REFERENCE: DAS-105X
; CURRENT APPLICATION NUMBER: US/11/375,551
; CURRENT FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: US/10/753,901
; PRIOR FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US 60/441,717
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16
; LENGTH: 1016
; TYPE: PRF
; ORGANISM: Xenorhabdus nematophilus
US-11-375-551-16

Query Match 1.5%; Score 101; DB 7; Length 1016;
Best Local Similarity 18.3%; Pred.No.53; Mismatches 229; Indels 410; Gaps 42;
Matches 162; Conservative 83

QY 41 KNPGLSGSLMTLSNV-----SCLS-----NTPARKITGDHLILILQKICPRLY--TGNP 86
Db 124 ESDNLPGRLLTITEQVKGENACITERLIWSGNTPAEK--GNN--LAGQCVVHYDPTGN 178
QY 87 TQACCSAKOLVLEASLSITKALLTRCPACSNFVNLHCHNTCSNQSLSFINVTRVAQIG 146
Db 179 QTNSSIS--LTSI--PLSITQQLK-----DDSEADWHGMD----- 209
QY 147 AQCLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATLAVGTWCMGVYGSALCNAQR----- 201
Db 210 -----ESGKNALAPESFTSVS-----TTDATGTVL-TSTDAAGNKRQIAYDV 251
QY 202 -----WLNFOQ-----DTGNGLAPLDITFHLLEP----- 225
Db 252 AGLQGSWLAALKQEQVIVKSLTYSAAQKLEEHNGI-----VTTYTEPETQVIGI 307
QY 226 -----GOAVGSGI-----OPLNEGVARCNESQ----- 247
Db 308 KTERPSGHAAGEKILQNLRYEYDVPVGNVLKSTNDABEITFRWNQKIVPENTVYDLSYOL 367
QY 248 ----GDDVATCSQDCAASCAPAIARQALDSTFYLG-----QMPGSLVLIILCSVPA 296
Db 368 VSVTGREMANIGRQKNQLPAL-----IDNNNTYNSRTYDYDRGNLTRI----- 414

QY 297 VTILLVGRVAPARDKSKMDVPPKGTSLSDKLSSTHTLLGQFQCGWGTWASWPLTIL 356
Db 415 -----RHNS-----PITGNNYTNMTVSDH----- 434
QY 357 VLSVIPVVALAAGLVFTELTTDP--VELWSAPNSQARSEKAFHQHGFPGFPRTNOVIL-- 412
Db 435 -----SNRAVLLELAQDPTQVDMLFTEGQHQLTRVPCQDLFWTPRDELOQQVILVN 484
QY 413 ---TAPNRSSRYDS-----LLIGPK-----NFSGLDLDLLE 443
Db 485 RENITPDQEFYHDADSQRVINKTHIQKGTNSQIORTLYLPELEWRTTYSG---NTLKE 540
QY 444 LLEL-----QERLRHLQ----- 455
Db 541 FLQVITVGSQQAQVRVLHWETGKPADISNDQLRYSYGNLIGSSGLELSDGQIISQEEY 600
QY 456 -----VMSPEAQRNISIQDIQYAPLNPDPNTSLYDCCINSLLQYFQNNRTLLLTANQT 508
Db 601 YPYGTAVWAARSQSEADYKTVRYSGKERDATGL-----YYGYRYQSWTGRWL 650
QY 509 LMGQTSQVDWKDHFLYCANAPLTFKDGDTALALSCMADYGAPVFPFLAIGYKGD-YSEA 567
Db 651 SVDPAGEVDGLNLFMRNRPVFSDDGR-----PPGQGVLAWIGKKAYRKA 698
QY 568 EALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVFTFAERSLEDEINR 627
Db 699 -----VNITTEHLEQGASFD-----TFL-----KLNK 721
QY 628 TTAEDLPIPATSYIVIFLYISLA-LGSYSSWSRVMVDSKATLGLGVAV---VLGAVMAA 683
Db 722 ----GLRTFVLGVGVASLGVAATAGASPM-----GIVGAALGGFVSGAV--- 763
QY 684 MGFF-----SYLGRSSSLVILQVVPFLVLSVGADNIF 715
Db 764 MGFFANNISEKIGEVLSYL-TRKRSVPVQVGAFVVTSLVTSALF 806

Search completed: April 11, 2006, 00:57:45
Job time : 106 secs

THIS PAGE IS BLANK

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE IS BLANK